

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2005, 00:40:07 ; Search time 3280 Seconds
(without alignments)
15334.303 Million cell updates/sec

Title: US-09-980-145-5
Perfect score: 1038
Sequence: 1 tcagaatgatgcctttg.....gtcatagattatratatc 1038

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1036.4	99.8	3501	6	AX259496 Sequence
2	1036.4	99.8	141028	6	AL513524 Human DNA
3	1032.4	99.5	1039	6	AX921196 Sequence
4	1032.4	99.5	1039	6	AX921198 Sequence
5	1032.4	99.5	1039	6	AX921206 Sequence
6	1028.4	99.1	1125	6	AX497139 Sequence
7	1021.4	98.4	1023	6	AX147844 Sequence
8	1021.4	98.4	1023	6	AX521893 Sequence
9	1020	98.3	1020	6	CO879788 Sequence
10	1020	98.3	1020	6	AX549353 Sequence
11	1020	98.3	1020	6	AF380185 Homo sapi
12	1018.4	98.1	1020	6	AX498188 Sequence
13	1018.4	98.1	1020	6	AF200627 Homo sapi
14	1018.4	98.1	1020	6	AF180374 Homo sapi
15	1015.4	97.8	1017	6	CO739824 Sequence
16	1015.4	97.8	1017	6	AX259494 Sequence
17	964.2	92.9	1017	9	AX153536 Macaca mu
18	869	83.7	1138	6	AX147842 Sequence
19	869	83.7	1138	6	AX521891 Sequence

20	866.4	83.5	933	6	AX657540 Sequence
21	694	66.9	962	6	AX147818 Sequence
22	694	66.9	962	6	AX521867 Sequence
23	660.8	63.7	245025	2	AC114062 Sequence
24	660.8	63.7	255076	2	AC128759 Rattus no
25	660.8	63.7	301304	2	AC130057 Rattus no
26	654.4	63.0	145990	10	AC117837 Rattus no
27	654.4	63.0	152612	2	AC118722 Mus muscu
28	653.4	62.9	999	10	AF380186 Rattus no
29	647	62.3	999	10	AF421352 Rattus no
30	643.8	62.0	999	10	AF380187 Mus muscu
31	637.4	61.4	999	6	AX497141 Sequence
32	381	36.7	613	11	BV063409 Sequence
33	305.8	29.5	173988	5	BX072572 Zebrafish
34	302.6	29.2	189341	10	AC117671 Mus muscu
35	302.6	29.2	214404	5	AL929507 Zebrafish
36	299.4	28.8	1044	10	AF380188 Rattus no
37	296	28.5	1114	9	AF112462 Homo sapi
38	295.6	28.5	2273	6	AR035954 Sequence
39	295.6	28.5	2273	6	AF112461 Homo sapi
40	292.8	28.2	1032	6	AX549293 Sequence
41	289.2	27.9	1049	9	HS088828 Homo sapien
42	279.8	27.0	1041	6	E13892 cDNA encodi
43	276.6	26.6	1041	9	AF112460 Homo sapi
44	270.8	26.1	919	9	BC067461 Homo sapi
45	269.2	25.9	919	9	BC067463 Homo sapi

ALIGNMENTS

RESULT 1	AX259496	3501 bp	DNA	linear	PAT 26-OCT-2001
LOCUS	AX259496				
DEFINITION	Sequence 3 from Patent WO0172841.				
ACCESSION	AX259496				
VERSION	AX259496.1	GI:16508654			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE					
AUTHORS	Kodira,C., Cravchik,A., di Francesco,V. and Beasley,E.M.				
TITLE	Isolated human G-protein coupled receptors that are members of the amineergic subfamily, nucleic acid molecules encoding human gpcr proteins, and uses thereof				
JOURNAL	Patent: WO 0172841-A 3 04-OCT-2001;				
PE Corporation (NY) (US)					
FEATURES					
source	1..3501				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
Query Match	99.8%	Score 1036.4;	DB 6;	Length 3501;	
Best Local Similarity	99.9%	Pred. No. 1.5e-232;			
Matches 1037;	Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;
1	TCAGGAATGATGCCCTTTGGCACATATTAATTAATTTCTGTGTGAAAAACAATGG	60			
1128	TCAGGAATGATGCCCTTTGGCACATATTAATTAATTTCTGTGTGAAAAACAATGG	1187			
61	TCAAATGATGTCGGTTCCTGTCACGTTAATGTCATATTAATTTGACCACTTC	120			
1188	TCAAATGATGTCGGTTCCTGTCACGTTAATGTCATATTAATTTGACCACTTC	1247			
121	GTGGCAATCGATGTTATTTCTATATACACTTCAACAACCTCATATCCCAACA	180			
1248	GTGGCAATCGATGTTATTTCTATATACACTTCAACAACCTCATATCCCAACA	1307			
181	AATGGCTCATTCATTCATGCGCACTGTGACCTTTCTTGGGGGTGTGTGTCATGCT	240			

```

Db      1308 AATTGGCTCAATTCATTCAGGACAGTGGACCTTCTCTGGGGGTGTGATGACGCT 1367
Qy      241 TACAGTATGGTGAATGCTGAGGACCTGTGGTATTTTGGAGAAAGCTCTGTAAT 300
Db      1368 TACAGTATGGTGAATGCTGAGGACCTGTGGTATTTTGGAGAAAGCTCTGTAAT 1427
Qy      301 CACACAAGCAGCAGCATTTATGCTGAGCTCAGCCTCCATTTTCAATTTGCTTCACTCC 360
Db      1428 CACACAAGCAGCAGCATTTATGCTGAGCTCAGCCTCCATTTTCAATTTGCTTCACTCC 1487
Qy      361 ATTGACCGGCTACATATGCTGTGTGTATCCATGAGATATTAAGCCAGATGAATATCTTG 420
Db      1488 ATTGACCGGCTACATATGCTGTGTGTATCCATGAGATATTAAGCCAGATGAATATCTTG 1547
Qy      421 GTTATTTGTGTATGATCTTCAATTTAGTGGAGTGTCCGCTGTTTTCATTTGGAAAG 480
Db      1548 GTTATTTGTGTATGATCTTCAATTTAGTGGAGTGTCCGCTGTTTTCATTTGGAAAG 1607
Qy      481 ATCTTTCTGAGAGCTTAACCTTCAAGGCGCTGAAGAGATATATTACAAATGTTCACTGC 540
Db      1608 ATCTTTCTGAGAGCTTAACCTTCAAGGCGCTGAAGAGATATATTACAAATGTTCACTGC 1667
Qy      541 AGAGAGAGTGTCTCTCTCTTTTACCAAAATATCTGGGGTACTGACCTTTATGACTTCT 600
Db      1668 AGAGAGAGTGTCTCTCTCTTTTACCAAAATATCTGGGGTACTGACCTTTATGACTTCT 1727
Qy      601 TTTTATATACCTGATCTATTTATGATGTATGCTATTTACAGAAATATCTTATCGCTAA 660
Db      1728 TTTTATATACCTGATCTATTTATGATGTATGCTATTTACAGAAATATCTTATCGCTAA 1787
Qy      661 GAACGAGCAAGATTAATTAAGTGCATTCAGAAAGCTCCAAATTGGATTGGAAATGAAA 720
Db      1788 GAACGAGCAAGATTAATTAAGTGCATTCAGAAAGCTCCAAATTGGATTGGAAATGAAA 1847
Qy      721 AATGGAATTTCAAAAGCAAGAAAGAAAGCTGGAAGACATTTGGGGATTTGTGTGGGA 780
Db      1848 AATGGAATTTCAAAAGCAAGAAAGAAAGCTGGAAGACATTTGGGGATTTGTGTGGGA 1907
Qy      781 GTTTTCTTAATATGCTGTGCTCTTTCTTATCTGTACAGTCAGTACGACCTTTTCTTAC 840
Db      1908 GTTTTCTTAATATGCTGTGCTCTTTCTTATCTGTACAGTACGACCTTTTCTTAC 1967
Qy      841 TACATTAATTCACCACTTATGATGATGTGTTGTTGGTCTTGAATCTTAC 900
Db      1968 TACATTAATTCACCACTTATGATGATGTGTTGTTGGTCTTGAATCTTAC 2027
Qy      901 TTTAATCCAAATGTTTATGCAATTTTCTATCCTTGTTTGAAGAAAGCACTGAAGATGATG 960
Db      2028 TTTAATCCAAATGTTTATGCAATTTTCTATCCTTGTTTGAAGAAAGCACTGAAGATGATG 2087
Qy      961 CTGTTTGGTAAATTTTCCAAAAGATTCATCCAGGTGTAATTAATTTTGGAAATGAGT 1020
Db      2088 CTGTTTGGTAAATTTTCCAAAAGATTCATCCAGGTGTAATTAATTTTGGAAATGAGT 2147
Qy      1021 TCATGAAATTAATTAAT 1038
Db      2148 TCATGAAATTAATTAAT 2165

```

```

RESULT 2
AL513524/c 141028 bp DNA linear PRI 11-MAY-2001
LOCUS      Human DNA sequence from clone RP11-295F4 on chromosome 6, complete
DEFINITION
ACCESSION  AL513524
VERSION     AL513524.8 GI:14031112
KEYWORDS   HTG.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens (human)
REFERENCE   1 (bases 1 to 141028)
            Bukacinski, Metaczo, Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

AUTHORS TITLE JOURNAL COMMENT

Clark G.
Direct Submission
Submitted (10-MAY-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On May 14, 2001 this sequence version replaced gi:1375152.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPep; Information on the WormPep
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chromosome6.html
RP11-295F4 is from the library RPCT-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/Dacpac/home.htm
VECTOR: DBAC3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-295F4. It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true left end of clone RP11-295F4 is at 1 in this sequence. The
true left end of clone RP1-55C23 is at 140929 in this sequence. The
true right end of clone RP11-80A23 is at 96540 in this sequence.

FEATURES
source
1..141028
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-295F4"
/clone_1b="RPCT-11.2"
52..267
/note="L2 repeat: matches 1314..1547 of consensus"
97..1029
/note="L2 repeat: matches 2669..2701 of consensus"
1493..1562
/note="2 copies 35 mer 95% conserved"
2086..2215
/note="MER94 repeat: matches 1..133 of consensus"
2441..2610
/note="L1MD3 repeat: matches 7578..7738 of consensus"
2616..2704
/note="L1MC3 repeat: matches 6701..6788 of consensus"
2717..2832
/note="L1uB repeat: matches 3..124 of consensus"
2833..3295
/note="L1uB2 repeat: matches 1..444 of consensus"
3296..3497
/note="L1uB repeat: matches 124..304 of consensus"
3809..3844
/note="L1MC/D repeat: matches 5568..5603 of consensus"
3881..4045
/note="L1MC4 repeat: matches 6460..6611 of consensus"
4046..4111
/note="MADE1 repeat: matches 1..80 of consensus"
4112..4885
/note="L1MC4 repeat: matches 5600..6460 of consensus"
5473..5887

/note="LIMC4 repeat: matches 5201. .5630 of consensus"
repeat_region 5888. .6069
/note="PAM repeat: matches 1. .173 of consensus"
repeat_region 6070. .7026
/note="LIMC4 repeat: matches 4226. .5201 of consensus"
repeat_region 7454. .9534
/note="LIMC4 repeat: matches 2110. .4239 of consensus"
repeat_region 9835. .9963
/note="LIMC4 repeat: matches 1968. .2101 of consensus"
repeat_region 10234. .10331
/note="LTR16A repeat: matches 59. .153 of consensus"
repeat_region 10731. .11121
/note="MER57B repeat: matches 1. .403 of consensus"
repeat_region 11122. .15589
/note="MER57-internal repeat: matches 1. .4612 of consensus"
repeat_region 15590. .15899
/note="AluY repeat: matches 1. .306 of consensus"
repeat_region 15900. .16478
/note="MER57-internal repeat: matches 4612. .5175 of consensus"
repeat_region 16480. .16587
/note="AluSg/x repeat: matches 200. .307 of consensus"
repeat_region 17166. .17532
/note="THE1B repeat: matches 1. .364 of consensus"
repeat_region 17728. .17925
/note="L2 repeat: matches 1636. .1836 of consensus"
repeat_region 18077. .18479
/note="THE1B-INTERNAL repeat: matches 1186. .1574 of consensus"
repeat_region 18485. .18845
/note="THE1B repeat: matches 1. .364 of consensus"
repeat_region 18846. .18885
/note="L1PA15 repeat: matches 6115. .6157 of consensus"
repeat_region 19862. .19979
/note="LIMC4 repeat: matches 7857. .7977 of consensus"
repeat_region 20016. .21137
/note="MER11A repeat: matches 1. .1266 of consensus"
repeat_region 21138. .21336
/note="LIMC4 repeat: matches 7654. .7844 of consensus"
repeat_region 21337. .21608
/note="AluY repeat: matches 39. .293 of consensus"
repeat_region 21609. .22560
/note="LIMC4 repeat: matches 6650. .7654 of consensus"
repeat_region 23089. .23122
/note="L2 copies 2 mer ac 82% conserved"
repeat_region 28852. .28920
/note="MIR repeat: matches 8. .77 of consensus"
repeat_region 29765. .29800
/note="L18 copies 2 mer ca 97% conserved"
repeat_region 30361. .30361
/note="Sequence from overlapping clone ba80a23 (AL589691). Assembly confirmed by restriction digest."
repeat_region 30587. .30722
/note="HERV1 repeat: matches 3314. .3453 of consensus"
repeat_region 31082. .31233
/note="HERV1 repeat: matches 2705. .2858 of consensus"
repeat_region 31761. .31969
/note="MLT1H repeat: matches 29. .294 of consensus"
repeat_region 32170. .32723
/note="MER67C repeat: matches 1. .506 of consensus"
repeat_region 33639. .33922
/note="AluX repeat: matches 3. .286 of consensus"
repeat_region 34841. .35352
/note="MLT1D repeat: matches 1. .505 of consensus"
repeat_region 35113. .35279
/note="Sequence from overlapping clone ba80a23 (AL589691). Assembly confirmed by restriction digest"
repeat_region 35703. .36065
/note="L2 repeat: matches 1141. .1515 of consensus"
repeat_region 36191. .36282
/note="46 copies 2 mer tc 62% conserved"
repeat_region 36321. .36350

/note="L5 copies 2 mer tc 86% conserved"
repeat_region 36450. .36652
/note="L2 repeat: matches 1675. .1914 of consensus"
repeat_region 37308. .38171
/note="MER45C repeat: matches 1. .895 of consensus"
repeat_region 40257. .40551
/note="AluB repeat: matches 20. .313 of consensus"
repeat_region 40619. .40973
/note="MLT1B repeat: matches 1. .390 of consensus"
repeat_region 42469. .43278
/note="L1P52 repeat: matches 5315. .6154 of consensus"
repeat_region 44183. .44258
/note="L1M4 repeat: matches 2886. .2965 of consensus"
repeat_region 44917. .45185
/note="AluSg repeat: matches 1. .295 of consensus"
repeat_region 46431. .46800
/note="L1PA13 repeat: matches 5787. .6156 of consensus"
repeat_region 46819. .46945
/note="HERV1 repeat: matches 397. .1283 of consensus"
repeat_region 46959. .47269
/note="AluB repeat: matches 1. .125 of consensus"
repeat_region 49450. .50350
/note="HERV1 repeat: matches 397. .1283 of consensus"
repeat_region 50724. .50941
/note="L1M4 repeat: matches 2184. .2413 of consensus"
repeat_region 50956. .51095
/note="L4 copies 35 mer 73% conserved"
repeat_region 50985. .51032
/note="L2 copies 4 mer tata 79% conserved"
repeat_region 51015. .51122
/note="L4 copies 27 mer 72% conserved"
repeat_region 51358. .51503
/note="L1M4 repeat: matches 2422. .2571 of consensus"
repeat_region 51491. .51668
/note="L1M1 repeat: matches 4563. .4733 of consensus"
repeat_region 51675. .51852
/note="AluSg repeat: matches 1. .179 of consensus"
repeat_region 51853. .52151
/note="L1M1 repeat: matches 4894. .5201 of consensus"
repeat_region 52194. .52911
/note="L1M1 repeat: matches 5284. .6017 of consensus"
repeat_region 54539. .54764
/note="L2 repeat: matches 2514. .2750 of consensus"

Query Match 99.8%; Score 1036.4; DB 9; Length 141028;
Best Local Similarity 99.9%; Pred. No. 1e-222; Indels 0; Gaps 0;
Matches 1037; Conservative 1; Indels 0; Gaps 0;

1 TCAGGATGATGCCCTTTGGCCACATATATTAATTTCTGTGTGAAAAACAATGG 60
132175 TCAGGATGATGCCCTTTGGCCACATATATTAATTTCTGTGTGAAAAACAATGG 132116

61 TCAATGATGTCGGTTCCTGTCAGTTTAAATGTCCTAATTTTGACACACATC 120
132115 TCAATGATGTCGGTTCCTGTCAGTTTAAATGTCCTAATTTTGACACACATC 132056

121 GTTGCATGATGATGATTTATTTCTATATGACATTCACCACTTATATCCCAACA 180
132056 GTTGCATGATGATGATTTATTTCTATATGACATTCACCACTTATATCCCAACA 131996

181 AATTGGCATATTCATTCATGCGCATGTCGATTCCTTCTTGGGGTGTCTGTCATGCT 240
131995 AATTGGCATATTCATTCATGCGCATGTCGATTCCTTCTTGGGGTGTCTGTCATGCT 131936

241 TACAGTATGATGATGATTCGCTGACAGCTGTGATTTTGGAGAAATCTTCTGTAATTT 300
131935 TACAGTATGATGATGATTCGCTGACAGCTGTGATTTTGGAGAAATCTTCTGTAATTT 131876

301 CACACAGACCGCATATATGTCGAGCTCAGCTCCATTTTTCATTTGCTTCAATCTCC 360
131875 CACACAGACCGCATATATGTCGAGCTCAGCTCCATTTTTCATTTGCTTCAATCTCC 131816

361 ATTGACCGCTACTATGCTGTGTGTATCATCATGAGATATTAAGCAAGATGATATCTTG 420

Db 131815 ATTACCGCTACTATGCTGTGTGATGCACTGAGATATTAAGCCAAAGAAATATCTTG 131756
Qy 421 GTTATTTGTGATGATCTTATTAAGTTGAGAGTCCCTGCTGTTTTGCAATTTGGAATG 480
Db 131755 GTTATTTGTGATGATCTTATTAAGTTGAGAGTCCCTGCTGTTTTGCAATTTGGAATG 131696
Qy 481 ATCTTTCGAGCTTAACCTCAAGAGCGCTGAGAGATATATTAACAACATGTTCACTGC 540
Db 131695 ATCTTTCGAGCTTAACCTCAAGAGCGCTGAGAGATATATTAACAACATGTTCACTGC 131636
Qy 541 AGAGAGAGTTCCTGCTGCTCTTTTGAAGAAATATCTGGGCTACTGACCTTTATGACTTCT 600
Db 131635 AGAGAGAGTTCCTGCTGCTCTTTTGAAGAAATATCTGGGCTACTGACCTTTATGACTTCT 131576
Qy 601 TTTTATATACCTGATCTATATATATATATGCTATATGCTATATACAGATATATCTATGCTTAA 660
Db 131575 TTTTATATACCTGATCTATATATATATATGCTATATGCTATATACAGATATATCTATGCTTAA 131516
Qy 661 GAAACAGGCAAGATTAATTAATGATGCAATCAGAGCTCCAAATTTGGAATGAAATGAA 720
Db 131515 GAAACAGGCAAGATTAATTAATGATGCAATCAGAGCTCCAAATTTGGAATGAAATGAA 131456
Qy 721 AATGGAATTTCAAGAGCAAGAAAGAAAGCTGTGAGACATTTGGGGAATTTGATGGGA 780
Db 131455 AATGGAATTTCAAGAGCAAGAAAGAAAGCTGTGAGACATTTGGGGAATTTGATGGGA 131396
Qy 781 GTTTTCTATATATGCTGTGCTGCTCTTCTTATCTGTACAGTCAATGAGCCCTTTCTTCA 840
Db 131395 GTTTTCTATATATGCTGTGCTGCTCTTCTTATCTGTACAGTCAATGAGCCCTTTCTTCA 131336
Qy 841 TACATATATCCACCTTACTTGAATGATGATGATTTGGTTGGCTACTTGAACCTTCA 900
Db 131335 TACATATATCCACCTTACTTGAATGATGATGATTTGGTTGGCTACTTGAACCTTCA 131276
Qy 901 TTTTATCCATGATGTTATGATCTTCTATCTGCTGTTTGAAGAACCTGAAAGATG 960
Db 131275 TTTTATCCATGATGTTATGATCTTCTATCTGCTGTTTGAAGAACCTGAAAGATG 131216
Qy 961 CTGTTTGGTAAATTTTCCAAAAGATTCATCAGGTGTAAATTTATTTTGGAAATGGAGT 1020
Db 131215 CTGTTTGGTAAATTTTCCAAAAGATTCATCAGGTGTAAATTTATTTTGGAAATGGAGT 131156
Qy 1021 TCATAGAAATTAATTAAT 1038
Db 131155 TCATAGAAATTAATTAAT 131138

RESULT 3
AX921196 1039 bp DNA linear PAT 18-DEC-2003
LOCUS AX921196
DEFINITION Sequence 189 from Patent WO02068652.
ACCESSION AX921196
VERSION AX921196.1 GI:40214830
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 Nov-x proteins and nucleic acids encoding same
JOURNAL Patent: WO 02068652-A 189 06-SEP-2002;
FEATURES
LOCATION/Qualifiers
1. 1039
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.5%; Score 1032.4; DB 6; Length 1039;
Best Local Similarity 99.9%; Pred. No. 1.5e-231;
Matches 1033; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 GAATGATGCCCTTTTGCCCAATATATTAATTTCCGTGTGAAAAACAACCTGTCGA 64
Db 1 GAATGATGCCCTTTTGCCCAATATATTAATTAATTTCCGTGTGAAAAACAACCTGTCGA 60
Qy 65 ATGATGTCGTGCTTCCCTGTCACAGTTTAATGTCCTCATTAATTCGACACACTGTTG 124
Db 61 ATGATGTCGTGCTTCCCTGTCACAGTTTAATGTCCTCATTAATTCGACACACTGTTG 120
Qy 125 GCATCTGATAGTATGTTGTTCTATATACACTTCAACACTTCAATCCCAACAAATT 184
Db 121 GCATCTGATAGTATGTTGTTCTATATACACTTCAACACTTCAATCCCAACAAATT 180
Qy 185 GGCCTATTCATTCATGCGCACTGTGAATCTTCTGCGGGGTCTGTGTCATGCTTACA 244
Db 181 GGCCTATTCATTCATGCGCACTGTGAATCTTCTGCGGGGTCTGTGTCATGCTTACA 240
Qy 245 GTATGATGATATCTGCTGAGACACTGTGTGATTTTGGAGAACTCTTGTAATTCACA 304
Db 241 GTATGATGATATCTGCTGAGACACTGTGTGATTTTGGAGAACTCTTGTAATTCACA 300
Qy 305 CAAGCACCACATTAATGCTGAGCTCAGCTTCAATTTTCCATTTGCTTTGATCCTCAATG 364
Db 301 CAAGCACCACATTAATGCTGAGCTCAGCTTCAATTTTCCATTTGCTTTGATCCTCAATG 360
Qy 365 ACCGCTACTATGCTGTGTGTGATCCTAGATATTAAGCCAAAGATATCTTGTTA 424
Db 361 ACCGCTACTATGCTGTGTGTGATCCTAGATATTAAGCCAAAGATATCTTGTTA 420
Qy 425 TTTGTGATGATATCTTATAGTTGAGTGTCCCTGCTGTTTTGCAATTTGGAATGATCT 484
Db 421 TTTGTGATGATATCTTATAGTTGAGTGTCCCTGCTGTTTTGCAATTTGGAATGATCT 480
Qy 485 TTCTGAGCTTAACCTCAAGAGCGCTGAGAGATATTAACAACATGTTCACTGAGAG 544
Db 481 TTCTGAGCTTAACCTCAAGAGCGCTGAGAGATATTAACAACATGTTCACTGAGAG 540
Qy 545 GAGGTCTGCTGCTCTTTTGAAGAAATATCTGGGGTATCTGACCTTTATGACTTTT 604
Db 541 GAGGTCTGCTGCTCTTTTGAAGAAATATCTGGGGTATCTGACCTTTATGACTTTT 600
Qy 605 ATATACCTGATCTATATGATATGTCATATACAGAAATATCTTATGCTTAAGAAC 664
Db 601 ATATACCTGATCTATATGATATGTCATATACAGAAATATCTTATGCTTAAGAAC 660
Qy 665 AGGCAAGATTAATTAATGATGCCATCAGAGCTCCAAATTTGATGGAATGAAATG 724
Db 661 AGGCAAGATTAATTAATGATGCCATCAGAGCTCCAAATTTGATGGAATGAAATG 720
Qy 725 GAATTTCAAGAGCAAGAAAGAAAGCTGTGAGACATTTGGGGATTTGATGGAGTT 784
Db 721 GAATTTCAAGAGCAAGAAAGAAAGCTGTGAGACATTTGGGGATTTGATGGAGTT 780
Qy 785 TCCATATATGCTGTGCTCTTCTTATATGTCAGTCATGAGACCTTTTCTTCACTACA 844
Db 781 TCCATATATGCTGTGCTCTTCTTATATGTCAGTCATGAGACCTTTTCTTCACTACA 840
Qy 845 TTATTCACCTACTTTGAATGATGTCATTTTGGTTGGCTACTTGAACCTTCACTTTA 904
Db 841 TTATTCACCTACTTTGAATGATGTCATTTTGGTTGGCTACTTGAACCTTCACTTTA 900
Qy 905 ATCCAAATGTTTATGATTTTCTATCTGTTTGAAGAAAGCACTGAAAGATGATGCTGT 964
Db 901 ATCCAAATGTTTATGATTTTCTATCTGTTTGAAGAAAGCACTGAAAGATGATGCTGT 960
Qy 965 TTGTTAAATTTTCCAAAAGATTCATCAGGTGTAAATTTATTTTGGAAATGAGTTGAT 1024
Db 961 TTGTTAAATTTTCCAAAAGATTCATCAGGTGTAAATTTATTTTGGAAATGAGTTGAT 1020
Qy 1025 AGAATTAATTAATTT 1038
Db 1021 AGAATTAATTAATTT 1034


```
RESULT 4
AX921198      1039 bp   DNA      linear   PAT 18-DEC-2003
LOCUS         Sequence 191 from Patent WO02068652.
ACCESSION     AX921198
VERSION       AX921198.1 GI:40214831
KEYWORDS
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS       Nov-x proteins and nucleic acids encoding same
TITLE         Patent: WO 02068652-A 191 06-SEP-2002;
JOURNAL       Location/Qualifiers
FEATURES
             source
             1..1039
             /organism="Homo sapiens"
             /mol_type="unassigned DNA"
             /db_xref="taxon:9606"

ORIGIN
Query Match      99.5%; Score 1032.4; DB 6; Length 1039;
Best Local Similarity 99.9%; Pred. No. 1.5e-231;
Matches 1033; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      5 GAATGATGCCCTTTTGGCCCAATATATATATTTCTCTGTGGAACCAACTGCTCA 64
DB      1 GAATGATGCCCTTTTGGCCCAATATATATATTTCTCTGTGGAACCAACTGCTCA 60
QY      65 ATGATGTCGGTCTTCCCTGTCAGTTTATGATGTCATATTTGACACACCTGCTG 124
DB      61 ATGATGTCGGTCTTCCCTGTCAGTTTATGATGTCATATTTGACACACCTGCTG 120
QY      125 GCAATCTGATGTTATTTGTTATATATACACTTCAACCAACTCATACCCCAAAAT 184
DB      121 GCAATCTGATGTTATTTGTTATATATACACTTCAACCAACTCATACCCCAAAAT 180
QY      185 GGCTCATTCATTCATGCGCACTGTCAGTCTTCTTCTGGGGTCTGTCATGCGCTTCA 244
DB      181 GGCTCATTCATTCATGCGCACTGTCAGTCTTCTTCTGGGGTCTGTCATGCGCTTCA 240
QY      245 GTATGATGATCTGCTGACGACTGTGGTATTTTGGAGAAGCTTCTGTAAATTCACA 304
DB      241 GTATGATGATCTGCTGACGACTGTGGTATTTTGGAGAAGCTTCTGTAAATTCACA 300
QY      305 CAAGCACCGACATTAATGCTGAGCTCAGCTCCATTTTCCATTTGCTTCATCTCAT 364
DB      301 CAAGCACCGACATTAATGCTGAGCTCAGCTCCATTTTCCATTTGCTTCATCTCAT 360
QY      365 ACCGCTACTATGCTGTGTGATCCACTGATATTAAGCCAAAGATGATATCTTGTTA 424
DB      361 ACCGCTACTATGCTGTGTGATCCACTGATATTAAGCCAAAGATGATATCTTGTTA 420
QY      425 TTGTGTGATGATCTTCAATTAAGTGGAGTGTCCCTGCTGTTTGGATTTGGATGATCT 484
DB      421 TTGTGTGATGATCTTCAATTAAGTGGAGTGTCCCTGCTGTTTGGATTTGGATGATCT 480
QY      485 TTCTGAGCTAAACTTCAAGGCGCTGAAGAGATATATTAACAACATGTTCACTGAGAG 544
DB      481 TTCTGAGCTAAACTTCAAGGCGCTGAAGAGATATATTAACAACATGTTCACTGAGAG 540
QY      545 GAGGTGCTCTGCTCTTCTTAGCAAAATATCTGGGGTACTGACCTTAAAGACTTCTTTT 604
DB      541 GAGGTGCTCTGCTCTTCTTAGCAAAATATCTGGGGTACTGACCTTAAAGACTTCTTTT 600
QY      605 ATATACCTGATCTATATGATATGATGCTATATTAAGATATATCTTATCGCTAAAGAC 664
DB      601 ATATACCTGATCTATATGATATGATGCTATATTAAGATATATCTTATCGCTAAAGAC 660
QY      665 AGGCAAGATTAATTAAGTATGCAATCAGAAGCTCCAAATGGAGTTGAAATGAAGAAATG 724
DB      661 AGGCAAGATTAATTAAGTATGCAATCAGAAGCTCCAAATGGAGTTGAAATGAAGAAATG 720
```

```
DB      661 AGGCAAGATTAATTAAGTATGCAATCAGAAGCTCCAAATGGAGTTGAAATGAAGAAATG 720
QY      725 GAATTTCACAAGCAAGAAAGAAAGAGCTGTGAGACATTTGGGGATTTGATGGAGTTT 784
DB      721 GAATTTCACAAGCAAGAAAGAAAGAGCTGTGAGACATTTGGGGATTTGATGGAGTTT 780
QY      785 TCTTAATATGCTGTGCTCTTCTTCTTATCTGTACAGTCATGGAACCTTTCTTCACTACA 844
DB      781 TCTTAATATGCTGTGCTCTTCTTCTTATCTGTACAGTCATGGAACCTTTCTTCACTACA 840
QY      845 TTATTTCCACTTACTTGAATGATGCTGTGATTTGGTTGGCTACTTGAACCTTACATTTA 904
DB      841 TTATTTCCACTTACTTGAATGATGCTGTGATTTGGTTGGCTACTTGAACCTTACATTTA 900
QY      905 ATCCAAATGTTATGATATTTTCTATCTCTGTTTGAAGAAAGCACTGAAGATGATGCTGT 964
DB      901 ATCCAAATGTTATGATATTTTCTATCTCTGTTTGAAGAAAGCACTGAAGATGATGCTGT 960
QY      965 TTGTTAAATTTTCCAAAGAAATTCATCAGGTGTAATTTATTTTGGAAATGAGTTTAT 1024
DB      961 TTGTTAAATTTTCCAAAGAAATTCATCAGGTGTAATTTATTTTGGAAATGAGTTTAT 1020
QY      1025 AGAATTATTAATTT 1038
DB      1021 AGAATTATTAATTT 1034

RESULT 5
AX921206      1039 bp   DNA      linear   PAT 18-DEC-2003
LOCUS         Sequence 199 from Patent WO02068652.
ACCESSION     AX921206
VERSION       AX921206.1 GI:40214835
KEYWORDS
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS       Nov-x proteins and nucleic acids encoding same
TITLE         Patent: WO 02068652-A 199 06-SEP-2002;
JOURNAL       Location/Qualifiers
FEATURES
             source
             1..1039
             /organism="Homo sapiens"
             /mol_type="unassigned DNA"
             /db_xref="taxon:9606"

ORIGIN
Query Match      99.5%; Score 1032.4; DB 6; Length 1039;
Best Local Similarity 99.9%; Pred. No. 1.5e-231;
Matches 1033; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      5 GAATGATGCCCTTTTGGCCCAATATATATATTTCTCTGTGGAACCAACTGCTCA 64
DB      1 GAATGATGCCCTTTTGGCCCAATATATATATTTCTCTGTGGAACCAACTGCTCA 60
QY      65 ATGATGTCGGTCTTCCCTGTCAGTTTATGATGTCATATTTGACACACCTGCTG 124
DB      61 ATGATGTCGGTCTTCCCTGTCAGTTTATGATGTCATATTTGACACACCTGCTG 120
QY      125 GCAATCTGATGTTATTTGTTATATATACACTTCAACCAACTCATACCCCAAAAT 184
DB      121 GCAATCTGATGTTATTTGTTATATATACACTTCAACCAACTCATACCCCAAAAT 180
QY      185 GGCTCATTCATTCATGCGCACTGTCAGTCTTCTTCTGGGGTCTGTCATGCGCTTCA 244
DB      181 GGCTCATTCATTCATGCGCACTGTCAGTCTTCTTCTGGGGTCTGTCATGCGCTTCA 240
QY      245 GTATGATGATCTGCTGACGACTGTGGTATTTTGGAGAAGCTTCTGTAAATTCACA 304
DB      241 GTATGATGATCTGCTGACGACTGTGGTATTTTGGAGAAGCTTCTGTAAATTCACA 300
```

OY		305	CAAGACCGGACAATTATGTCTGAGCTCAGCCCTCATTTTTCATTTTGTCCTTGCATCTCCATTG	367
Db		301	CAGACACCGGACATTATAGCTGAGCTCAGCCCTCATTTTTCATTTTGTCCTTGCATCTCCATTG	367
OY		365	ACCGETACTANGCTGTGTGTATCCATCGAGATATAAAGGCAAGAATGAATCTTTGGTTA	422
Db		361	ACCGETACTANGCTGTGTGTATCCATCGAGATATAAAGGCAAGAATGAATCTTTGGTTA	422
OY		425	TTTGNGNATGATCTTCACTTAGTGGAGTGTCCCTGCTGTTTTTGGCATTTGGAAATGATCT	484
Db		421	TTTGNGNATGATCTTCACTTAGTGGAGTGTCCCTGCTGTTTTTGGCATTTGGAAATGATCT	484
OY		485	TTTCGAGCTTAACCTTCAAAGGCGCTGAAGAGATATATTAACAACATGTTCACTGCAAG	544
Db		481	TTTCGAGCTTAACCTTCAAAGGCGCTGAAGAGATATATTAACAACATGTTCACTGCAAG	544
OY		545	GAGGTGTCTGTCTTCTTTAGCAAAAATATCTGGGATAGTACACTTATGACTTCTTTT	604
Db		541	GAGGTGTCTGTCTTCTTTAGCAAAAATATCTGGGATAGTACACTTATGACTTCTTTT	600
OY		605	ATATACCTGGATCTCATATATGTTANGTCTTATTAACAATATCTTATGCTTAAGAAC	664
Db		601	ATATACCTGGATCTCATATATGTTANGTCTTATTAACAATATCTTATGCTTAAGAAC	660
OY		665	AGGCAAGATTAAATTAGTAGTCCAAATCAGAAGCTCCAAATTGGATTGGAAAAATG	724
Db		661	AGGCAAGATTAAATTAGTAGTCCAAATCAGAAGCTCCAAATTGGATTGGAAAAATG	720
OY		725	GAATTTCAACAACCAAGAAAGAGAAAGCTGTGAAGCATTTGGGATTTGATGGAGTTT	784
Db		721	GAATTTCAACAACCAAGAAAGAGAAAGCTGTGAAGCATTTGGGATTTGATGGAGTTT	780
OY		785	TCTTAATATGCTGTGTCCTCTTTTATCTGTACAGTCANTGAGCCCTTTTCTTCACTACA	844
Db		781	TCTTAATATGCTGTGTCCTCTTTTATCTGTACAGTCANTGAGCCCTTTTCTTCACTACA	840
OY		845	TTATTTCCACCTACCTTGAATGATGTGTGATTTGTTGGCTACCTTGAACCTCACATTTA	904
Db		841	TTATTTCCACCTACCTTGAATGATGTGTGATTTGTTGGCTACCTTGAACCTCACATTTA	900
OY		905	ATCCAATGTTATGCAATTTTCTATCCCTGGTTTAAAGAAACACTGAAGATGATGCTGT	964
Db		901	ATCCAATGTTATGCAATTTTCTATCCCTGGTTTAAAGAAACACTGAAGATGATGCTGT	960
OY		965	TTGGTAAATTTTCCAAAAGATTCAATCCAGSTGTAAATATTTTGGATTTGAGTTGAT	1022
Db		961	TTGGTAAATTTTCCAAAAGATTCAATCCAGSTGTAAATATTTTGGATTTGAGTTGAT	1020
OY		1025	AGATTATTAATATTT 1038	
Db		1021	AGATTATTAATATTT 1034	
RESULT 6				
LOCUS	AX497139			
DEFINITION	Sequence 1 from Patent WO022801.	1125 bp	DNA	linear PAT 26-SEP-2002
ACCESSION	AK497139			
VERSION	AK497139.1	GI:23342534		
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;			
JOURNAL	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE	Bunzow,J.R., Grandy,D.K. and Sondere,M. Mammalian receptor genes and uses Patent: WO 022801-A 1 21-MAR-2002; Oregon Health & Science University (US) Location/Qualifiers			
FEATURES				
source	1..1125	/organism="Homo sapiens"		
		/mol_type="unassigned DNA"		

CDS
 /db_xref="taxon:9606"
 21..1040
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD48641.1"
 /db_xref="GI:23342535"
 /translation="MMPFCH3N"
 IVSIHFKQLHTPTWMLIHSMATVDFILGCTVMPVSRSAHECHYREBEVCKHTHS
 DIMLSASFIFHLSFISIDRYVAADPLRKYAKMNILVYCWVIFISMVPAFAGMIR
 EQRRLISDANQQLQIGLEKKNISQSKERKAVKVLGIWGFLLCMCFPLCTVMDPEP
 LHSIDLPITNDVLIFWGYLNSFENMVAFFYPWFRKRLKRWLFGLKIFQKDSRCKLE
 LRLSS"
 ORIGIN
 Query Match 99.1%; Score 1028.4; DB 6; Length 1125;
 Best Local Similarity 99.4%; Pred. No. 1,3e-230;
 Matches 1032; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 TCAGGAATGATGCCCCCTTTTGGCCAAATPAAATTAATATATCTGTGTGAATAACAATCG 60
 Db 15 TCAGGAATGATGCCCTTTTGGCCAAATPAAATTAATATATCTGTGTGAATAACAATCG 74
 QY 61 TCAGGAATGATGCCGCTTCCCTGTACAGCTTAAATGCTGTCAATATTCGACCACTC 120
 Db 75 TCAGGAATGATGCCGCTTCCCTGTACAGCTTAAATGCTGTCAATATTCGACCACTC 134
 QY 121 GTTGGCAATCTGAATGATTATTGTTCTATATACACTTCAACCAACTTCATACCCCA 180
 Db 135 GTTGGCAATCTGAATGATTATTGTTCTATATACACTTCAACCAACTTCATACCCCA 194
 QY 181 AATTGGCTCATTCATTCATGCGCATGCGCACTGTGGACTTCTCTGGGGTGTGTCAATGCT 240
 Db 195 AATTGGCTCATTCATTCATGCGCATGCGCACTGTGGACTTCTCTGGGGTGTGTCAATGCT 254
 QY 241 TACAGTATGAGAGATCTGCTGAGACACTGTTGGATTTTGGAGAAGCTCTGTGAAAT 300
 Db 255 TACAGTATGAGAGATCTGCTGAGACACTGTTGGATTTTGGAGAAGCTCTGTGAAAT 314
 QY 301 CACACAAACACCGACATTAATGCTGAGCTCAAGCTCCATTTTCCATTTGTTTCATCTCC 360
 Db 315 CACACAAACACCGACATTAATGCTGAGCTCAAGCTCCATTTTCCATTTGTTTCATCTCC 374
 QY 361 AATGACCGCTATCTATGCTGTGTGAGACCACTGAGATATAAAGCAAGATTAATCTTG 420
 Db 375 AATGACCGCTATCTATGCTGTGTGAGACCACTGAGATATAAAGCAAGATTAATCTTG 434
 QY 421 GTTAATTTGTGATGATCTTCATTAAGTTGAAGTGTCCGCTGTTTGTGCAATTTGGAA 480
 Db 435 GTTAATTTGTGATGATCTTCATTAAGTTGAAGTGTCCGCTGTTTGTGCAATTTGGAA 494
 QY 481 ATCTTTTCGAGAGCTAACTTCAAAAGCGCTGAAGAGATATATTACAACATGTTCACTGC 540
 Db 495 ATCTTTTCGAGAGCTAACTTCAAAAGCGCTGAAGAGATATATTACAACATGTTCACTGC 554
 QY 541 AGAGAGAGTGTCTGTCTCTTTAGCAAAATATCTGGGGTACTGACCTTAATGACTTCT 600
 Db 555 AGAGAGAGTGTCTGTCTCTTTAGCAAAATATCTGGGGTACTGACCTTAATGACTTCT 614
 QY 601 TTTTATATACCTGAGATCTATATGTTATGTCTATTAAGAAATAATCTTAATGCTTAA 660
 Db 615 TTTTATATACCTGAGATCTATATGTTATGTCTATTAAGAAATATCTTAATGCTTAA 674
 QY 661 GAACAGGCAAGATTATTTAGTAGTGCCATTCGAGAGCTCCAAATGGATTTGAAATGAA 720
 Db 675 GAACAGGCAAGATTATTTAGTAGTGCCATTCGAGAGCTCCAAATGGATTTGAAATGAA 734
 QY 721 AATGAATTTCAAAAGCAAAAGAAAGAAAGCTGTGAAGACAATGGAGATTTGTATGGA 780
 Db 735 AATGAATTTCAAAAGCAAAAGAAAGAAAGCTGTGAAGACAATGGAGATTTGTATGGA 794
 QY 781 GTTTTCTTAATATGCTGTGCCCCCTTTCTTTATCTGTACAGTCAATGACCTTTTCTTAC 840

Db 795 GTTTCCTATATAGCTGTGCCCCCTTCTTATCTGTACAGTCAGACCCCTTTCCTAC 854
Qy 841 TACATATATCCACCTACTTGTGAATGATGTTGTTGGCTACTTGAACCTACA 900
Db 855 TCAATATATCCACCTACTTGTGAATGATGTTGTTGGCTACTTGAACCTACA 914
Qy 901 TTTAATCCAAATGTTATGATCTTCTTCTGTTTGAAGACATGAAATGATG 960
Db 915 TTTAATCCAAATGTTATGATCTTCTTCTGTTTGAAGACATGAAATGATG 974
Qy 961 CTGTTGTGTAATAATTTCCAAAAAGATCATCAGGTGTAATTTATTTTGAATGAT 1020
Db 975 CTGTTGTGTAATAATTTCCAAAAAGATCATCAGGTGTAATTTATTTTGAATGAT 1034
Qy 1021 TCATAGAAATTTATATTT 1038
Db 1035 TCATAGAAATTTATATTT 1052

RESULT 7
AX147844 1023 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 89 from Patent WO0136473.
ACCESSION AX147844
VERSION AX147844.1 GI:14346843
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1
Vogeli, G., Wood, L.S., Parodi, L.A., Hiesch, R.R., Lind, P.,
Slichter, J., Schellin, K.A., Kayes, P.S., Bannigan, C.M., Ruff, V.,
Seifert, T., and Huff, R.M.
Novel 9 protein-coupled receptors
Patent: WO 0136473-A 89 25-MAY-2001;
JOURNAL PHARMACIA & UPJOHN COMPANY (US)
FEATURES
source 1..1023
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 98.4%; Score 1021.4; DB 6; Length 1023;
Best Local Similarity 99.9%; Pred. No. 5.7e-229;
Matches 1022; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 361 GACCGCTACTATGCTGTGTGATCCTGAGATATAAGCCAAATGATATCTTGGT 420
Qy 424 ATTGTGTGATGATCTTCAATAGTTGAGATGCTGCTGCTTGTGCAATTTGGAATGATC 483
Db 421 ATTGTGTGATGATCTTCAATAGTTGAGATGCTGCTGCTTGTGCAATTTGGAATGATC 480
Qy 484 TTTCTGAGCTTAACTTCAAGAGCCCTGAGAGATATATTCAACATGTTCACTGACA 543
Db 481 TTTCTGAGCTTAACTTCAAGAGCCCTGAGAGATATATTCAACATGTTCACTGACA 540
Qy 544 GGAGGTGCTGTCTTCTTGAAGAAATATCTGAGGATCTGACCTTTATGACTTCTTT 603
Db 541 GGAGGTGCTGTCTTCTTGAAGAAATATCTGAGGATCTGACCTTTATGACTTCTTT 600
Qy 604 TATATACCTGATCTATATAGTTATGATGCTTATACAAATATCTTATGCTTAAAGA 663
Db 601 TATATACCTGATCTATATAGTTATGATGCTTATACAAATATCTTATGCTTAAAGA 660
Qy 664 CAGCAAGATTAATAGTATGATGCTCAATCAGAGCTCAAAATTTGAAATGAAAAAT 723
Db 661 CAGCAAGATTAATAGTATGATGCTCAATCAGAGCTCAAAATTTGAAATGAAAAAT 720
Qy 724 GGAATTTCACAAAAGCAAAAGAAAGAAAGCTGGAAGCAATTTGGGATTTGTGAGAGTT 783
Db 721 GGAATTTCACAAAAGCAAAAGAAAGAAAGCTGGAAGCAATTTGGGATTTGTGAGAGTT 780
Qy 784 TTCTTAATATGCTGCTGCTTCTTATATCTGATCAAGCAATGACCTTTCTTCACTAC 843
Db 781 TTCTTAATATGCTGCTGCTTCTTATATCTGATCAAGCAATGACCTTTCTTCACTAC 840
Qy 844 ATTATTTCAACCTACTTGAATGATGTTGTTGTTGCTACTTGAATCTTCACTAT 903
Db 841 ATTATTTCAACCTACTTGAATGATGTTGTTGTTGCTACTTGAATCTTCACTAT 900
Qy 904 AATCCAAATGTTTATGATCTTCTTATCTGTTTGAAGAAAGCACTGAAGATGATCTG 963
Db 901 AATCCAAATGTTTATGATCTTCTTATCTGTTTGAAGAAAGCACTGAAGATGATCTG 960
Qy 964 TTTGTGTAATAATTTCCAAAAAGATCATCAGGTGTAATTTATTTTGAATGAGTTCA 1023
Db 961 TTTGTGTAATAATTTCCAAAAAGATCATCAGGTGTAATTTATTTTGAATGAGTTCA 1020
Qy 1024 TAG 1026
Db 1021 TAG 1023

RESULT 8
AX521893 1023 bp DNA linear PAT 24-OCT-2002
LOCUS AX521893
DEFINITION Sequence 89 from Patent WO02064789.
ACCESSION AX521893
VERSION AX521893.1 GI:24410795
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1
Lind, P., Parodi, L.A., Vogeli, G., and Wood, L.S.
G protein-coupled receptor
Patent: WO 02064789-A 89 22-OCT-2002;
JOURNAL PHARMACIA & UPJOHN COMPANY (US)
FEATURES
source 1..1023
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 98.4%; Score 1021.4; DB 6; Length 1023;
Best Local Similarity 99.9%; Pred. No. 5.7e-229;

Matches 1022; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY	4	GGAAATGACCCCTTTGGCCAAATATATATATTTCTGTGTGAAAAACAATGGTCA	63
Db	1	GGAAATGATCCCTTTGGCCAAATATATATATTTCTGTGTGAAAAACAATGGTCA	66
OY	64	AATGATGTCGGCTTCCTGTACAGTTTAATGAGTCTATATATCTGACACATCGTT	12
Db	61	AATGATGTCGGTCTTCCTGTACAGTTTAATGAGTCTATATATCTGACACATCGTT	12
OY	124	GGCAATCTGATAGTATATGTTTCTATATCAACTTCAACAACTTCATACCCCAAAAT	18
Db	121	GGCAATCTGATAGTATATGTTTCTATATCAACTTCAACAACTTCATACCCCAAAAT	18
OY	184	TGGCTCATTTCCATGGCCACTGTGACCTTTCTTGAGGAGTCTGATAGCTTAC	24
Db	181	TGGCTCATTTCCATGGCCACTGTGACCTTTCTTGAGGAGTCTGATAGCTTAC	24
OY	244	AGTATGAGAGATCTGCTGAGACATCTGTTGATTTTGGAGAAGTCTGTATAAATTCAC	30
Db	241	AGTATGAGAGATCTGCTGAGACATCTGTTGATTTTGGAGAAGTCTGTATAAATTCAC	30
OY	304	ACAAGCACCGACATTATGCTGAGCTCAGCCTCCATTTCCATTTGTCCTTATCTCCATT	36
Db	301	ACAAGCACCGACATTATGCTGAGCTCAGCCTCCATTTCCATTTGTCCTTATCTCCATT	36
OY	364	GACCGCTACTATGCTGTGTGATCTCACTGATATTAAGCCAAAGATGAATATCTGGTT	42
Db	361	GACCGCTACTATGCTGTGTGATCTCACTGATATTAAGCCAAAGATGAATATCTGGTT	42
OY	424	ATTGTGTGATGATCTTCATAGTGTGAGTGTCCCTGTGTTTGGATTTTGGAAATGATC	48
Db	421	ATTGTGTGATGATCTTCATAGTGTGAGTGTCCCTGTGTTTGGATTTTGGAAATGATC	48
OY	484	TTTCTGAGCTAAACCTCAAAAGGCGCTGAAGAGATATATTAACAACATGTTCACTGAGA	54
Db	481	TTTCTGAGCTAAACCTCAAAAGGCGCTGAAGAGATATATTAACAACATGTTCACTGAGA	54
OY	544	GGAAGTGTCTGTCTCTTCTTAGCAAAATATCTGGGGTATCTGACCTTATAGACTCTTTT	60
Db	541	GGAAGTGTCTGTCTCTTCTTAGCAAAATATCTGGGGTATCTGACCTTATAGACTCTTTT	60
OY	604	TATATGCTGGATCTATATATGTTATGTGTCTATTAACAATATATCTTATCGCTAAAGAA	66
Db	601	TATATGCTGGATCTATATATGTTATGTGTCTATTAACAATATATCTTATCGCTAAAGAA	66
OY	664	CAGGCAAGATTAATTAATGATGCAATCAGAACTCCAAATTTGGAATGGAATGAAT	72
Db	661	CAGGCAAGATTAATTAATGATGCAATCAGAACTCCAAATTTGGAATGGAATGAAT	72
OY	724	GGAAATTCACAAAGCAAAAGAAAGAAAGCTGTGAGACATTTGGGGATTTGATGGAGTT	78
Db	721	GGAAATTCACAAAGCAAAAGAAAGAAAGCTGTGAGACATTTGGGGATTTGATGGAGTT	78
OY	784	TTTCTATATATGCTGTGCTTCCTTATCTGTACAGTATGGAACCTTTCCTTCACTAC	84
Db	781	TTTCTATATATGCTGTGCTTCCTTATCTGTACAGTATGGAACCTTTCCTTCACTAC	84
OY	844	ATTATTTCAACCTTAAGTATGTTTCTTATCTGATTTGGCTTGTGCTTGAACCTTACATTT	90
Db	841	ATTATTTCAACCTTAAGTATGTTTCTTATCTGATTTGGCTTGTGCTTGAACCTTACATTT	90
OY	904	AATCCAAATGTTTATGCAATTTTCTATCTGGTATTAGAAAAGCACTGAAGATGATGCTG	96
Db	901	AATCCAAATGTTTATGCAATTTTCTATCTGGTATTAGAAAAGCACTGAAGATGATGCTG	96
OY	964	TTTGGTAAATTTTCCAAAAGATTCATCCAGGTGTAATATTTTGGAAATGAGTTCA	102
Db	961	TTTGGTAAATTTTCCAAAAGATTCATCCAGGTGTAATATTTTGGAAATGAGTTCA	102
OY	1024	TAG	1026
Db	1021	TAG	1023

RESULT 9			
CO879788			
LOCUS	CO879788	1020 bp	DNA
DEFINITION	Sequence 1 from Patent WO2004083851.		linear
ACCESSION	CO879788		
VERSION	CO879788.1	GI:54033726	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE			
AUTHORS	1 Golz, S., Brueggemeier, U. and Geerts, A.		
TITLE	Diagnostics and therapeutics for diseases associated with trace		
	amine receptor 1 (cal)		
	Patent: WO 2004083851-A 1 30-SEP-2004;		
JOURNAL	Bayer Healthcare AG (DE)		
FEATURES			
Source	Location/Qualifiers		
	1..1020		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		

Query Match	98.3%	Score 1020;	DB 6;	Length 1020;
Best Local Similarity	100.0%;	Pred. No. 1.2e-228;		
Matches 1020;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps
QY 7	ATGATGCCCCCTTTTGGCCACAAATATAATTAATTAATTTCTGTGTGAAAAAACAATGTCAAAT	666		
DB 1	ATGATGCCCCCTTTTGGCCACAAATATAATTAATTAATTTCTGTGTGAAAAAACAATGTCAAAT	600		
QY 67	GATGTCGGGCTTCCCTGTACAGTTTAAGTGTCTAATTTCTGACACACTCTGTGGC	12		
DB 61	GATGTCGGGCTTCCCTGTACAGTTTAAGTGTCTAATTTCTGACACACTCTGTGGC	12		
QY 127	AATTCGATGATTAATTTGTTCTAATATCAACTTCAACACTTCATACCCCAACAATTTGG	18		
DB 121	AATTCGATGATTAATTTGTTCTAATATCAACTTCAACACTTCATACCCCAACAATTTGG	18		
QY 187	CTCATTCATTCATAGGCCACCTGTGSACTTCTCTGGGGGTCTGTCAATGCTTACGT	24		
DB 181	CTCATTCATTCATAGGCCACCTGTGSACTTCTCTGGGGGTCTGTCAATGCTTACGT	24		
QY 247	ATGTGTGAGATCTGCTGAGCACTGTGTGTAATTTGGAGAAATCTTCTGTAAATTCACACA	30		
DB 241	ATGTGTGAGATCTGCTGAGCACTGTGTGTAATTTGGAGAAATCTTCTGTAAATTCACACA	30		
QY 307	AGCACCGACATTAATGTGAGACTCAGCCTCCATTTTCCATTGTCCTTCAATCTCCATTGAC	36		
DB 301	AGCACCGACATTAATGTGAGACTCAGCCTCCATTTTCCATTGTCCTTCAATCTCCATTGAC	36		
QY 367	CGCATCAATATGCTGTGTGTATTCACATCGAATATAAAGCCAAAGATGAATATCTGGTATTT	42		
DB 361	CGCATCAATATGCTGTGTGTATTCACATCGAATATAAAGCCAAAGATGAATATCTGGTATTT	42		
QY 427	TGTGTGATGATCTTCAATTTAGTTGSAATGTCCCGCTGTTTTTGGATTGGAAATGATCTTT	48		
DB 421	TGTGTGATGATCTTCAATTTAGTTGSAATGTCCCGCTGTTTTTGGATTGGAAATGATCTTT	48		
QY 487	CTGAGAGTAACTTCAAAAGCGCTGAGAGATATAATTAACAAACATGTTCACTGACAGAGA	54		
DB 481	CTGAGAGTAACTTCAAAAGCGCTGAGAGATATAATTAACAAACATGTTCACTGACAGAGA	54		
QY 547	GATGTCTGTCTTCTTTAGCAAAATATCTGGGGAGTACGACTTATAGACTTCTTTTAT	60		
DB 541	GATGTCTGTCTTCTTTAGCAAAATATCTGGGGAGTACGACTTATAGACTTCTTTTAT	60		
QY 607	ATACCTGGAATTAATTAATTAATGTGTCTAATTAACAAGATATAATCTTATGCTAAAGAACAG	66		
DB 601	ATACCTGGAATTAATTAATTAATGTGTCTAATTAACAAGATATAATCTTATGCTAAAGAACAG	66		

QY 667 GCAGATTAAATAGTATGATCCATCAGAAAGCTCCAAATTGGATTGGAATGAAAAATGGA 726
DB 661 GCAGATTAAATAGTATGATCCATCAGAAAGCTCCAAATTGGATTGGAATGAAAAATGGA 720
QY 727 ATTTACAAAGCAAGAAAGAAAGAAAGCTGTGAAGACATTGGGGATTGTGATGGAGATTTC 786
DB 721 ATTTACAAAGCAAGAAAGAAAGAAAGCTGTGAAGACATTGGGGATTGTGATGGAGATTTC 780
QY 787 CTAAATAGTGTGTCCTTTCTTTATCTGTACAGTCAGACCTTTTCTTACATCACT 846
DB 781 CTAAATAGTGTGTCCTTTCTTTATCTGTACAGTCAGACCTTTTCTTACATCACT 840
QY 847 ATTCACCTACTTGAATGATGTGTGATTTGGTGGCTACTTGAACCTTCACTTTAT 906
DB 841 ATTCACCTACTTGAATGATGTGTGATTTGGTGGCTACTTGAACCTTCACTTTAT 900
QY 907 CCAATGTTTATGACATTTTCTATCTGTTTGAAGAAAGCACTGAAGATGATGCTTT 966
DB 901 CCAATGTTTATGACATTTTCTATCTGTTTGAAGAAAGCACTGAAGATGATGCTTT 960
QY 967 GGTAAATTTTCCAAAAGATTCATCCAGGTGTAAATTTTGGAAATGAGTCAATG 1026
DB 961 GGTAAATTTTCCAAAAGATTCATCCAGGTGTAAATTTTGGAAATGAGTCAATG 1020

RESULT 10
AX549353
LOCUS AX549353 1020 bp DNA linear PAT 26-NOV-2002
DEFINITION Sequence 638 from Patent WO2061087.
ACCESSION AX549353
VERSION AX549353.1 GI:25813987
KEYWORDS

SOURCE
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS Homo sapiens
TITLE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1

Burner, G.C., Roush, C.L. and Brown, J.P.

Antigenic peptides, such as for G protein-coupled receptors
(GPCRs), antibodies thereto, and systems for identifying such
antigenic peptides

Patent: WO 02061087-A 638 08-AUG-2002;

LifeSpan Biosciences, Inc. (US)

location/Qualifiers

FEATURES

source

1.1020

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 98.3%; Score 1020; DB 6; Length 1020;

Best Local Similarity 100.0%; Pred. No. 1.2e-228;

Matches 1020; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ATGATGCCCTTTGGCACAATATATATATTTCTGTGTAAGAAACAACTGTGCAAT 66

DB 1 ATGATGCCCTTTGGCACAATATATATATTTCTGTGTAAGAAACAACTGTGCAAT 60

QY 67 GATGTCGGTCTTCCCTGTACAGTTAATGATGTCATATTTCTGACACACTGCTGGC 126

DB 61 GATGTCGGTCTTCCCTGTACAGTTAATGATGTCATATTTCTGACACACTGCTGGC 120

QY 127 AATCTGATAGTATTTGTTCTATATCACACTTCAACAACTTCATACCCCAAAATGG 186

DB 121 AATCTGATAGTATTTGTTCTATATCACACTTCAACAACTTCATACCCCAAAATGG 180

QY 187 CTCATTCAATTCAGGCACTGAGCACTTCTTGAGGGTGTCTGGTCACTGCTTACAGT 246

DB 181 CTCATTCAATTCAGGCACTGAGCACTTCTTGAGGGTGTCTGGTCACTGCTTACAGT 240

QY 247 ATGTGAGATCTGCTGAGCACTGTGGATTTTGGAAAGTCTTCTGTAATTCACACA 306

|||||

DB 241 ATGTGAGATCTGCTGAGCACTGTGGATTTTGGAAAGTCTTCTGTAATTCACACA 300
QY 307 AGCACCGCATTTATGCTGAGCTCAGCCCTCCATTTTCCATTTGTCTTTCATCCATGAC 366
DB 301 AGCACCGCATTTATGCTGAGCTCAGCCCTCCATTTTCCATTTGTCTTTCATCCATGAC 360
QY 367 CGCTACTATGCTGTGTGTGATCCTGAGATATTAAGCAAGATGATATCTTGGTTAT 426
DB 361 CGCTACTATGCTGTGTGTGATCCTGAGATATTAAGCAAGATGATATCTTGGTTAT 420
QY 427 TGTGTGATGATCTTCAATGATGAGATGCTCCCTGCTGTTTGGATTTGGATATCTTT 486
DB 421 TGTGTGATGATCTTCAATGATGAGATGCTCCCTGCTGTTTGGATTTGGATATCTTT 480
QY 487 CTGAGCTTAACTTCAAGAGGCTGAGAGATATTTCAACATGTTCACTGACAGAGA 546
DB 481 CTGAGCTTAACTTCAAGAGGCTGAGAGATATTTCAACATGTTCACTGACAGAGA 540
QY 547 GGTGCTCTGCTCTTCTTTAGCAAAATATCTGGGCTACTGACCTTTATGACTTTTAT 606
DB 541 GGTGCTCTGCTCTTCTTTAGCAAAATATCTGGGCTACTGACCTTTATGACTTTTAT 600
QY 607 ATACCTGATCTATTTATGATATGATGCTTATTAACAATATATCTTATGCTTAAAGACAG 666
DB 601 ATACCTGATCTATTTATGATATGATGCTTATTAACAATATATCTTATGCTTAAAGACAG 660
QY 667 GCAGATTAAATAGTATGATCCATCAGAAAGCTCCAAATTGGATTGGAATGAAAAATGGA 726
DB 661 GCAGATTAAATAGTATGATCCATCAGAAAGCTCCAAATTGGATTGGAATGAAAAATGGA 720
QY 727 ATTTACAAAGCAAGAAAGAAAGAAAGCTGTGAAGACATTGGGGATTGTGATGGAGATTTC 786
DB 721 ATTTACAAAGCAAGAAAGAAAGAAAGCTGTGAAGACATTGGGGATTGTGATGGAGATTTC 780
QY 787 CTAAATAGTGTGTCCTTTCTTTATCTGTACAGTCAGACCTTTTCTTACATCACT 846
DB 781 CTAAATAGTGTGTCCTTTCTTTATCTGTACAGTCAGACCTTTTCTTACATCACT 840
QY 847 ATTCACCTACTTGAATGATGTGTGATTTGGTGGCTTGAACCTTCACTTCACTTAT 906
DB 841 ATTCACCTACTTGAATGATGTGTGATTTGGTGGCTTGAACCTTCACTTCACTTAT 900
QY 907 CCAATGTTTATGACATTTTCTATCTGTTTGAAGAAAGCACTGAAGATGATGCTTT 966
DB 901 CCAATGTTTATGACATTTTCTATCTGTTTGAAGAAAGCACTGAAGATGATGCTTT 960
QY 967 GGTAAATTTTCCAAAAGATTCATCCAGGTGTAAATTTTGGAAATGAGTCAATG 1026
DB 961 GGTAAATTTTCCAAAAGATTCATCCAGGTGTAAATTTTGGAAATGAGTCAATG 1020

RESULT 11

AF380185 1020 bp DNA linear PRI 01-AUG-2001

LOCUS AF380185

DEFINITION Homo sapiens trace amine receptor 1 (TA1) gene, complete cde.

ACCESSION AF380185

VERSION AF380185.1 GI:14600073

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

AUTHORS

Borowczyk, B., Adam, N., Jones, K.A., Radetz, R., Artyushyn, R.,

Ogozalek, K.L., Durkin, M.M., Lakhiani, P.P., Bonini, J.A.,

Pachizana, S., Boyle, N., Pu, X., Kouranov, E., Lichblau, H.,

Ochoa, F.Y., Branchek, T.A. and Gerald, C.

Trace amines: identification of a family of mammalian G

protein-coupled receptors

Proc. Natl. Acad. Sci. U.S.A. 98 (16), 8966-8971 (2001)

JOURNAL MEDLINE

PUBMED 21374364

REFERENCE 11459929

2 (bases 1 to 1020)

AUTHORS Adham, N., Raddatz, R., Artyemyshyn, R., Durkin, M.M., Ogozalek, K.L., Lakhiani, P.P., Kouzanova, E.V., Ochoa, Y., Boyle, N., Pu, X., Branchek, T., Gerald, C., Blackburn, T.P., and Borowsky, B.

TITLE Pharmacological characterization of a G protein-coupled receptor, Tal, activated by trace amines and amphetamines: Evidence for major species differences in drug profile and tissue distribution

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 1020)

AUTHORS Borowsky, B., Ogozalek, K.L., Jones, K.A. and Adham, N.

TITLE Direct Submission

JOURNAL Submitted (12-MAY-2001) Target Discovery & Assessment, Synaptic Pharmaceutical Corp., 215 College Road, Paramus, NY 07652, USA

FEATURES Location/Qualifiers

Source 1..1020

gene /organism="Homo sapiens"

mRNA /mol_type="genomic DNA"

CDS /db_xref="taxon:9606"

gene /map="6q23.2"

mRNA <1..>1020

CDS /gene="TA1"

gene <1..>1020

CDS /product="TA1"

gene /product="trace amine receptor 1"

gene 1..1020

gene /gene="TA1"

gene /note="G protein-coupled receptor"

gene /product="trace amine receptor 1"

gene /protein_id="AAK71236.1"

gene /db_xref="GI:1460074"

gene /translation="MMPCHNINISCVNNNSNDVRASLYSLMVLILITLVGNLIVISLHFKQIAHTPTWMLHSMATVDLGLCLVMPYSMAASHCHWPCVPECKHISTDMLSSASIFHLFSIDRYAVCDPLRYKAMNITLVCVMFISMSPYAVAFEMITLBNFKAIEIYYKHVHCRCGCGSVFSKISGLTFTMSFVYDGSIMLCVYVRIYIAKBOARLISDANOKLOIGLEKNGISOSKSKRAKVTGIWGFPLICWCPFFICTVMDPFLHVIPTPLNDVLIMFGYILNSTFENPMTAFYPMFRKALKNMLFKKIFQKDSRCKLLELSS"

ORIGIN

Query Match 98.3%; Score 1020; DB 9; Length 1020;

Best Local Similarity 100.0%; Pred. No. 1.2e-228;

Matches 1020; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

7 ATGATGCCCTTTGGCCACAATATTAATATTTCTGTGTAAGAAACAAGTGTCAAT 66

1 ATGATGCCCTTTGGCCACAATATTAATATTTCTGTGTAAGAAACAAGTGTCAAT 60

67 GATGTCGGCTTCCCTGTGACGTTAATGAGTCTCATTAATTCGACCACTGCTGGC 126

61 GATGTCGGCTTCCCTGTGACGTTAATGAGTCTCATTAATTCGACCACTGCTGGC 120

127 AATCTGATAGTATTTGTTCTATATCACTTCAAAACAATTCATACCCCAAAATGG 186

121 AATCTGATAGTATTTGTTCTATATCACTTCAAAACAATTCATACCCCAAAATGG 180

187 CTGATTCATTCATGCGCACTGTGCACTTTCTTGCGGGGTCTGTGATGCTTACAGT 246

181 CTGATTCATTCATGCGCACTGTGCACTTTCTTGCGGGGTCTGTGATGCTTACAGT 240

247 ATGGTGAATGTCGCGGACGACGTTGATATTTGGAGAAGCTTCTGMAAATTCACACA 306

241 ATGGTGAATGTCGCGGACGACGTTGATATTTGGAGAAGCTTCTGMAAATTCACACA 300

307 AGCACCGACATTATGCTGAGCTCAGCTCATATTTTCCATTTGTCTTCATCTCCATTGAC 366

301 AGCACCGACATTATGCTGAGCTCAGCTCAGCTCATATTTTCCATTTGTCTTCATCTCCATTGAC 360

367 CGCTACTATGCTGTGTGATGCACTGAGATATAAGCCAAAGATGAATATCTTGTTAT 426

361 CGCTACTATGCTGTGTGATGCACTGAGATATAAGCCAAAGATGAATATCTTGTTAT 420

427 TGTGTGATGATCTTCATTAATTGAGAGTGTCCTGCTGTTTGTGATTTGGAATGATCTTT 486

AUTHORS Adham, N., Raddatz, R., Artyemyshyn, R., Durkin, M.M., Ogozalek, K.L., Lakhiani, P.P., Kouzanova, E.V., Ochoa, Y., Boyle, N., Pu, X., Branchek, T., Gerald, C., Blackburn, T.P., and Borowsky, B.

TITLE Pharmacological characterization of a G protein-coupled receptor, Tal, activated by trace amines and amphetamines: Evidence for major species differences in drug profile and tissue distribution

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 1020)

AUTHORS Borowsky, B., Ogozalek, K.L., Jones, K.A. and Adham, N.

TITLE Direct Submission

JOURNAL Submitted (12-MAY-2001) Target Discovery & Assessment, Synaptic Pharmaceutical Corp., 215 College Road, Paramus, NY 07652, USA

FEATURES Location/Qualifiers

Source 1..1020

gene /organism="Homo sapiens"

mRNA /mol_type="genomic DNA"

CDS /db_xref="taxon:9606"

gene /map="6q23.2"

mRNA <1..>1020

CDS /gene="TA1"

gene <1..>1020

CDS /product="TA1"

gene /product="trace amine receptor 1"

gene 1..1020

gene /gene="TA1"

gene /note="G protein-coupled receptor"

gene /product="trace amine receptor 1"

gene /protein_id="AAK71236.1"

gene /db_xref="GI:1460074"

gene /translation="MMPCHNINISCVNNNSNDVRASLYSLMVLILITLVGNLIVISLHFKQIAHTPTWMLHSMATVDLGLCLVMPYSMAASHCHWPCVPECKHISTDMLSSASIFHLFSIDRYAVCDPLRYKAMNITLVCVMFISMSPYAVAFEMITLBNFKAIEIYYKHVHCRCGCGSVFSKISGLTFTMSFVYDGSIMLCVYVRIYIAKBOARLISDANOKLOIGLEKNGISOSKSKRAKVTGIWGFPLICWCPFFICTVMDPFLHVIPTPLNDVLIMFGYILNSTFENPMTAFYPMFRKALKNMLFKKIFQKDSRCKLLELSS"

ORIGIN

Query Match 98.1%; Score 1018.4; DB 6; Length 1020;

Best Local Similarity 99.9%; Pred. No. 2.9e-228;

Matches 1019; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

7 ATGATGCCCTTTGGCCACAATATTAATATTTCTGTGTAAGAAACAAGTGTCAAT 66

1 ATGATGCCCTTTGGCCACAATATTAATATTTCTGTGTAAGAAACAAGTGTCAAT 60

67 GATGTCGGCTTCCCTGTGACGTTAATGAGTCTCATTAATTCGACCACTGCTGGC 126

Db 61 GATGTCGGGCTTCCCTGTACAGTTTAAAGGCTCATTAATCTGACCACTCGTTGGC 120

Qy 127 AATCGATGATATGTTTCTATATCACCTTCAAAACAATTATATCCCAACAATGG 186

Db 121 AATCGATATGTTATGTTTCTATATCACCTTCAAAACAATTATATCCCAACAATGG 180

Qy 187 CTCATTGATCCATGAGCCACCTGAGCACTTCTTCTGGGGTGTCTGTCATGCTTACAGT 246

Db 181 CTCATTGATCCATGAGCCACCTGAGCACTTCTTCTGGGGTGTCTGTCATGCTTACAGT 240

Qy 247 ATGGTGAATCTGCTGAGACACTGTTGGATTTTGGAGAAAGCTTCTGTAATAATCACACA 306

Db 241 ATGGTGAATCTGCTGAGACACTGTTGGATTTTGGAGAAAGCTTCTGTAATAATCACACA 300

Qy 307 AGCAGCGACATTAAGCTGAGCTCAGCCCTCAATTTTCCATTTTCTTCTGATCCATGAC 366

Db 301 AGCAGCGACATTAAGCTGAGCTCAGCCCTCAATTTTCCATTTTCTTCTGATCCATGAC 360

Qy 367 CGCTACTATGCTGTGTGATCCACTGAGATATTAAGCCAAAGATGAAATCTTGGTTAT 426

Db 361 CGCTACTATGCTGTGTGATCCACTGAGATATTAAGCCAAAGATGAAATCTTGGTTAT 420

Qy 427 TGTGTGATGATCTTCAATTAAGTTGAGAGTGTCCCTGCTGTTTGGCATTTGGAATGATCTTT 486

Db 421 TGTGTGATGATCTTCAATTAAGTTGAGAGTGTCCCTGCTGTTTGGCATTTGGAATGATCTTT 480

Qy 487 CTGGAGCTAACTTCAAAAGCCGCTGAAGATATTAACAACATGTTTCACTGAGAGAGA 546

Db 481 CTGGAGCTAACTTCAAAAGCCGCTGAAGATATTAACAACATGTTTCACTGAGAGAGA 540

Qy 547 GGTGCTCTGTCTCTTTAGCAAAATATCTGGGGTACTACCTTTATGACTCTTTTAT 606

Db 541 GGTGCTCTGTCTCTTTAGCAAAATATCTGGGGTACTACCTTTATGACTCTTTTAT 600

Qy 607 ATACTGATCTATTAATGTTATGTTGTCTATTAACAAGATATATCTTAAGCTTAAGAAGCAG 666

Db 601 ATACTGATCTATTAATGTTATGTTGTCTATTAACAAGATATATCTTAAGCTTAAGAAGCAG 660

Qy 667 GCAAAATTAATTAAGATGCCAATCAAGAACCTCCAAATTTGGAATGGAATAATGA 726

Db 661 GCAAAATTAATTAAGATGCCAATCAAGAACCTCCAAATTTGGAATGGAATAATGA 720

Qy 727 ATTTCACAAAGCAAAAGAAAGGCTGGAAGACATTTGGGATTTGTGAGGAGTTTC 786

Db 721 ATTTCACAAAGCAAAAGAAAGGCTGGAAGACATTTGGGATTTGTGAGGAGTTTC 780

Qy 787 CTAAATATGCTGTGCTCTTTCTTTATCTGTACAGTCATGAGACCTTTTCTTCACTACAT 846

Db 781 CTAAATATGCTGTGCTCTTTCTTTATCTGTACAGTCATGAGACCTTTTCTTCACTACAT 840

Qy 847 ATTTCACCTACTTTGAATGATGTGATTTGGTCTACTTGAACCTTCACTTAT 906

Db 841 ATTTCACCTACTTTGAATGATGTGATTTGGTCTACTTGAACCTTCACTTAT 900

Qy 907 CCAATGTTTATGACATTTTCTATCTTGGTTTGAAGAAAGCACTGAAGATGATGCTGTT 966

Db 901 CCAATGTTTATGACATTTTCTATCTTGGTTTGAAGAAAGCACTGAAGATGATGCTGTT 960

Qy 967 GGTAAATTTTCCAAAAGATTCATCAGGTGTAAATTAATTTTGGAAATGAGTTCAATG 1026

Db 961 GGTAAATTTTCCAAAAGATTCATCAGGTGTAAATTAATTTTGGAAATGAGTTCAATG 1020

RESULT 13

AF200627 1020 bp DNA linear PRI 10-JAN-2002

LOCUS Homo sapiens putative catecholamine receptor gene, complete cds.

DEFINITION AF200627

ACCESSION AF200627.1 GI:10441576

VERSION

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE

AUTHORS

Bunzow, J.R., Sonders, M.S., Artavanisoul, S., Harrison, L.M., Zhang, G., Quigley, D.I., Darland, T., Suchland, K.L., Pasumamula, S., Kennedy, J.L., Olson, S.B., Magenis, R.B., Amara, S.G., and Grandy, D.K. Amphetamine, 3,4-methylenedioxymethamphetamine, lysergic acid diethylamide, and metabolites of the catecholamine neurotransmitters are agonists of a rat trace amine receptor Mol. Pharmacol. 60 (6), 1181-1188 (2001)

JOURNAL

MEDLINE

21580235

11723224

2 (bases 1 to 1020)

REFERENCE

Bunzow, J.R. and Grandy, D.K.

AUTHORS

TITLE

Direct Submision

JOURNAL

Submitted (22-SEP-1999) Physiology and Pharmacology, Oregon Health Sciences University, 3181 S.W. Sam Jackson Park Rd., Portland, OR 97201, USA

FEATURES

source

1..1020

location/Qualifiers

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

<1..>1020

/product="putative catecholamine receptor"

1..1020

/codon_start=1

/product="putative catecholamine receptor"

/protein_id="AA017112.1"

/db_xref="gi:10441577"

/translation="MPECHNIINISCVKNNMSNDVRASLSYLVLIILTLVGNLIVIVSIHFQKLTPTMLIHSNATVDPLGCVLPVMSVRSABHCWFSEVFEKIHSTDIILNSASIFHLSPISIDRYAVCPDRKKNMLIVCNWIFISMSIPAYAFAMIFLEINPKAGEIYVGHVHRCGGSVPSFISVLPVMSFVYBGSIMLCVYRIVYIAKEQARLISPNQKLOIGLEMRKGISQSKRKAIVKILGIYMGVPLICMCFPICTWDPELHYIIPPLNDVLIIMFGYLNSTFENVYAFYIPWFRKALKMLFGRITQKSSRCKLPLELS"

ORIGIN

Query Match 98.1%; Score 1018.4; DB 9; Length 1020;

Best Local Similarity 99.9%; Pred. No. 2,9e-226;

Matches 1019; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 ATGATGCCCTTTGGCAAAATTAATTAATTTCTGTTGGAAGAAACAAGTGTCAAT 66

Db 1 ATGATGCCCTTTGGCAAAATTAATTAATTTCTGTTGGAAGAAACAAGTGTCAAT 60

Qy 67 GATGTCGGTGTCTCCCTGTACAGTTTAAAGTGTCTCATTAATTTGTGACACACTGCTGGC 126

Db 61 GATGTCGGTGTCTCCCTGTACAGTTTAAAGTGTCTCATTAATTTGTGACACACTGCTGGC 120

Qy 127 AATCGATGATATGTTTCTATATCACCTTCAAAACAATTATATCCCAACAATGG 186

Db 121 AATCGATGATATGTTTCTATATCACCTTCAAAACAATTATATCCCAACAATGG 180

Qy 187 CTCATTGATCCATGAGCCACCTGAGCACTTCTTCTGGGGTGTCTGTCATGCTTACAGT 246

Db 181 CTCATTGATCCATGAGCCACCTGAGCACTTCTTCTGGGGTGTCTGTCATGCTTACAGT 240

Qy 247 ATGGTGAATCTGCTGAGACACTGTTGGATTTTGGAGAAAGCTTCTGTAATAATCACACA 306

Db 241 ATGGTGAATCTGCTGAGACACTGTTGGATTTTGGAGAAAGCTTCTGTAATAATCACACA 300

Qy 307 AGCAGCGACATTAAGCTGAGCTCAGCCCTCAATTTTCCATTTTCTTCTGATCCATGAC 366

Db 301 AGCAGCGACATTAAGCTGAGCTCAGCCCTCAATTTTCCATTTTCTTCTGATCCATGAC 360

Qy 367 CGCTACTATGCTGTGTGATCCACTGAGATATTAAGCCAAAGATGAAATCTTGGTTAT 426

Db 361 CGCTACTATGCTGTGTGATCCACTGAGATATTAAGCCAAAGATGAAATCTTGGTTAT 420

Qy 427 TGTGTGATGATCTTCAATTAAGTTGAGAGTGTCCCTGCTGTTTGGCATTTGGAATGATCTTT 486

Db 421 TGTGTGATGATCTTATAGTGAAGTGTCCGTGCTTTTTCATTTGGAATGATCTTT 480
Qy 487 CTGAGAGCTAACTTTCAAGGCGCTGAAGATATATTACAAATGTTCACTGACAGGA 546
Db 481 CTGAGAGCTAACTTTCAAGGCGCTGAAGATATATTACAAATGTTCACTGACAGGA 540
Qy 547 GGTGGCTGTCTCTTTTGAAGCAAAATATCGGGTACTGACCTTTTGTGACTTTTAT 606
Db 541 GGTGGCTGTCTCTTTTGAAGCAAAATATCGGGTACTGACCTTTTGTGACTTTTAT 600
Qy 607 ATTCCTGATCTATTTATGTTATGTCTATACAGAAATATCTTATCGCTAAAGACAG 666
Db 601 ATACCTGATCTATTTATGTTATGTCTATACAGAAATATCTTATCGCTAAAGACAG 660
Qy 667 GCAGATTAATTAATGATGCAATGCAAGCTCAAAATTTGAAATGAAATGGA 726
Db 661 GCAGATTAATTAATGATGCAATGCAAGCTCAAAATTTGAAATGAAATGGA 720
Qy 727 ATTTCAAAAGCAAAAGAAAGAAAGCTGTGAAGACATTTGGGATTTGATGGAGTTTC 786
Db 721 ATTTCAAAAGCAAAAGAAAGAAAGCTGTGAAGACATTTGGGATTTGATGGAGTTTC 780
Qy 787 CTAAATATGCTGTGCCCCCTTTCTTTATCTGTACAGTATGAGCCCTTTCTTCACTACAT 846
Db 781 CTAAATATGCTGTGCCCCCTTTCTTTATCTGTACAGTATGAGCCCTTTCTTCACTACAT 840
Qy 847 ATTCACCTACTTGTGAATGATGTTGATTTGGTTTGGCTACTGAACTCTACATTTAT 906
Db 841 ATTCACCTACTTGTGAATGATGTTGATTTGGTTTGGCTACTGAACTCTACATTTAT 900
Qy 907 CCAATGTTTATGATCTTTTCTATCTTGTGTTAGAAAGCACTGAAGATGATGCTTTT 966
Db 901 CCAATGTTTATGATCTTTTCTATCTTGTGTTAGAAAGCACTGAAGATGATGCTTTT 960
Qy 967 GGTAAATTTTCCAAAAGATTCATCCAGGTGTAATTTTGTGAATGAGTTGATG 1026
Db 961 GGTAAATTTTCCAAAAGATTCATCCAGGTGTAATTTTGTGAATGAGTTGATG 1020

RESULT 14
AY180374
LOCUS AY180374 1020 bp mRNA linear PRI 19-JAN-2003
DEFINITION Homo sapiens trace amine receptor 1 mRNA, complete cds.
ACCESSION AY180374
VERSION AY180374.1 GI:27803881
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Kopatz,S.A., Aronstam,R.S. and Sharma,S.V.
TITLE Isolation of cDNA coding for Human Trace Amine Receptor (TAR) 1
JOURNAL Unpublished
AUTHORS Kopatz,S.A., Aronstam,R.S. and Sharma,S.V.
TITLE Direct Submission
JOURNAL Submitted (15-NOV-2002) Guthrie cDNA Resource Center, Guthrie
Research Institute, 1 Guthrie Square, Sayre, PA 18840, USA
FEATURES
SOURCE
1..1020
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..1020
/note="G-protein coupled receptor"
/codon_start=1
/product="trace amine receptor 1"
/protein_id="AA022154.1"
/db_xref="GI:27803882"
/translation="MMPCHNIIINISCVKNMSNDVRAISLYSLVLIILITIVGNLIV
IVSISHPKOLHTPTNMLIHMAIVDPLGLGLVMPYSVRAEHEHWYRGVEUCKHTST
DIMSSASIFHLSPISIDRYAVCDPLRYRAKKNILIVICMIFISMSVPAVFAFMIF

ORIGIN
Query Match 98.1%; Score 1018.4; DB 9; Length 1020;
Best Local Similarity 99.9%; Pred. No. 2.9e-228;
Matches 1019; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
LELNFKGAEIYYKHVHCRGGCSVFESKISGLVLFMTSFLYIPGSLMLCVYRIYLIAK
EQARLISDANKLQKLEMGNGISQSKSRKAVKTLGIWGVFLCMCEPFLCTVWDPF
LHYIIPPTLNDVLIMFGYLNSTFNDPMVYAFYPMWRKALKNMLFKIKIPOKSSRCKLF
LELS"

Qy 7 ATGATGCCCTTTTGGCAATATATATATATTTTCTGTGTGAAGAAACAGCTGCAAT 66
Db 1 ATGATGCCCTTTTGGCAATATATATATATTTCTGTGTGAAGAAACAGCTGCAAT 60
Qy 67 GATGTCGCTGCTTCCCTGTACAGTTTAATGAGTGTCTATTAATTTGACACACCTGGTGC 126
Db 61 GATGTCGCTGCTTCCCTGTACAGTTTAATGAGTGTCTATTAATTTGACACACCTGGTGC 120
Qy 127 AATCTGATAGTTATTTGTTCTTATATACACTTCAACAACTTCATACCCCAAAATTTG 186
Db 121 AATCTGATAGTTATTTGTTCTTATATACACTTCAACAACTTCATACCCCAAAATTTG 180
Qy 187 CTCATTCAATCCATGSCCACTGTGGACTTTCTCTGGGGTGTCTGTATCAGCTTACAGT 246
Db 181 CTCATTCAATCCATGSCCACTGTGGACTTTCTCTGGGGTGTCTGTATCAGCTTACAGT 240
Qy 247 ATGCTGATATGCTGTGACACTGTGTTGTTTGGAGAAAGCTTCTGTAAATTCACACA 306
Db 241 ATGCTGATATGCTGTGACACTGTGTTGTTTGGAGAAAGCTTCTGTAAATTCACACA 300
Qy 307 AGCACCGACATTATGCTGAGCTGAGCTGAGCTTTCATTTGCTTTCATCTTCATCCATGAC 366
Db 301 AGCACCGACATTATGCTGAGCTGAGCTGAGCTTTCATTTGCTTTCATCTTCATCCATGAC 360
Qy 367 CGCTACTATGCTGTGTGTGATCCACTGAGATATTAAGCCAAAGATGAATATCTTGTTAT 426
Db 361 CGCTACTATGCTGTGTGTGATCCACTGAGATATTAAGCCAAAGATGAATATCTTGTTAT 420
Qy 427 TGTGTGATGATCTTCAATTAATGAGTGTCCCTGCTGCTTTTGTGATTTGAAATGATCTTT 486
Db 421 TGTGTGATGATCTTCAATTAATGAGTGTCCCTGCTGCTTTTGTGATTTGAAATGATCTTT 480
Qy 487 CTGAGCTAAACCTTCAAGCGCTGAAGAGATATATTACAAACATGTTCACTGACAGGA 546
Db 481 CTGAGCTAAACCTTCAAGCGCTGAAGAGATATATTACAAACATGTTCACTGACAGGA 540
Qy 547 GGTGGCTGTCTCTTTTGAAGCAAAATATCGGGTACTGACCTTTTATGACTTCTTTAT 606
Db 541 GGTGGCTGTCTCTTTTGAAGCAAAATATCGGGTACTGACCTTTTATGACTTCTTTAT 600
Qy 607 ATACCTGATCTATTTATGTTATGTCTATACAGAAATATCTTATCGCTAAAGACAG 666
Db 601 ATACCTGATCTATTTATGTTATGTCTATACAGAAATATCTTATCGCTAAAGACAG 660
Qy 667 GCAGATTAATTAATGATGCAATGCAAGCTCAAAATTTGAAATGAAATGGA 726
Db 661 GCAGATTAATTAATGATGCAATGCAAGCTCAAAATTTGAAATGAAATGGA 720
Qy 727 ATTTCAAAAGCAAAAGAAAGAAAGCTGTGAAGACATTTGGGATTTGATGGAGTTTC 786
Db 721 ATTTCAAAAGCAAAAGAAAGAAAGCTGTGAAGACATTTGGGATTTGATGGAGTTTC 780
Qy 787 CTAAATATGCTGTGCCCCCTTTCTTTATCTGTACAGTATGAGCCCTTTCTTCACTACAT 846
Db 781 CTAAATATGCTGTGCCCCCTTTCTTTATCTGTACAGTATGAGCCCTTTCTTCACTACAT 840
Qy 847 ATTCACCTACTTGTGAATGATGTTGATTTGGCTACTGAACTCTACATTTAT 906
Db 841 ATTCACCTACTTGTGAATGATGTTGATTTGGCTACTGAACTCTACATTTAT 900
Qy 907 CCAATGTTTATGATCTTTTCTATCTTGTGTTAGAAAGCACTGAAGATGATGCTGTTT 966
Db 901 CCAATGTTTATGATCTTTTCTATCTTGTGTTAGAAAGCACTGAAGATGATGCTGTTT 960

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2005, 00:39:32 ; Search time 463 Seconds
(without alignments)
13271.472 Million cell updates/sec

Title: US-09-980-145-5
Perfect score: 1038
Sequence: 1 tcaggaatgacgaccttgc.....gtccatagatcatatcatt 1038

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: geneeqn19808:*
2: geneeqn19908:*
3: geneeqn20008:*
4: geneeqn20018:*
5: geneeqn20018:*
6: geneeqn20028:*
7: geneeqn20028:*
8: geneeqn20038:*
9: geneeqn20038:*
10: geneeqn20038:*
11: geneeqn20038:*
12: geneeqn20048:*
13: geneeqn20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1038	100.0	1038	4 AAC90247	AAC90247 Human SNO
2	1036.4	99.8	3501	4 AAH26920	AAH26920 Human G p
3	1032.4	99.5	1039	6 ADH48915	ADH48915 NOV84B co
4	1032.4	99.5	1039	6 ADH48905	ADH48905 NOV81A co
5	1032.4	99.5	1039	6 ADH48907	ADH48907 NOV81B co
6	1028.4	99.1	1125	6 AB158316	AB158316 Human tra
7	1021.4	98.4	1023	6 AAH51013	AAH51013 Human ncp
8	1021.4	98.4	1023	6 AB870246	AB870246 DNA encod
9	1020	98.3	1020	13 AB242578	AB242578 Human tra
10	1020	98.3	1020	13 AD873690	AD873690 Human tra
11	1018.4	98.1	1020	6 ABT04871	ABT04871 Human G p
12	1018.4	98.1	1020	12 AD030050	AD030050 Human GPC
13	1015.4	97.8	1017	4 AAH26919	AAH26919 Human G p
14	869	83.7	1138	6 AAH51012	AAH51012 Human ncp
15	869	83.7	1138	6 AB870245	AB870245 DNA encod
16	866.4	83.5	933	10 AD012791	AD012791 Human GPC
17	864	66.9	962	4 AAH51000	AAH51000 Human ncp
18	864	66.9	962	6 AB870233	AB870233 DNA encod
19	665.6	64.1	1101	4 AAC90246	AAC90246 Rat SNORF
20	653	62.9	1031	4 AAC90276	AAC90276 Mouse SNO

21	642.4	61.9	996	12 ADO30340	ADO30340 Mouse GPC
22	637.4	61.4	999	6 AB158317	AB158317 Rat trace
23	573	55.2	573	4 AAC90245	AAC90245 Human SNO
24	302.6	29.2	1041	12 ADO29146	ADO29146 Mouse nov
25	295.6	28.5	2273	2 AAH41998	AAH41998 Human HNH
26	292.8	28.2	1032	2 AB242548	AB242548 Human G p
27	292.8	28.2	1032	12 ADO29928	ADO29928 Human GPC
28	283	27.3	1029	12 ADO30218	ADO30218 Mouse GPC
29	279.8	27.0	1098	6 AB081178	AB081178 Human SNO
30	279.8	27.0	1098	10 ADE81114	ADE81114 Human SNO
31	276.6	26.6	1041	8 AB259359	AB259359 Human GPC
32	269.2	25.9	921	2 ABT94894	ABT94894 Human G p
33	269.2	25.9	921	10 ABV74592	ABV74592 Human G-p
34	269.2	25.9	1635	10 ADF70509	ADF70509 Orphan re
35	266	25.6	915	8 AB259360	AB259360 Human GPC
36	266	25.6	921	8 AB242547	AB242547 Human G p
37	266	25.6	921	12 ADO05721	ADO05721 Human G p
38	266	25.6	921	12 ADO29769	ADO29769 Human nov
39	263.2	25.4	921	12 ADO29075	ADO29075 Mouse nov
40	234.2	22.6	1044	12 ADO29149	ADO29149 Mouse nov
41	233	22.4	831	8 AB259361	AB259361 Mouse GPC
42	231.2	22.3	1034	8 AB269479	AB269479 Human G-p
43	231.2	22.3	1035	6 ABQ74527	ABQ74527 Human G p
44	231.2	22.3	1035	12 ADO30051	ADO30051 Human GPC
45	231.2	22.3	1038	6 ABT04870	ABT04870 Human G p

ALIGNMENTS

RESULT 1
AAC90247
ID AAC90247 standard; DNA; 1038 BP.
XX
XX AAC90247;
AC
XX 14-MAR-2001 (first entry)
DT
XX
XX Human SNORF3 receptor DNA.
DE
XX
XX SNORF3; inflammation; arthritis; neurological disorder; infection;
KW bone disease; respiratory disorder; asthma; cancer; cardiovascular; ds.
OS Homo sapiens.
XX
XX MO200073449-A1.
PN
XX
XX 07-DEC-2000.
PD
XX
XX 26-MAY-2000; 2000MO-US014654.
PF
XX
XX 28-MAY-1999; 99US-00322257.
PR 06-OCT-1999; 99US-00413433.
XX
XX (SYNA-) SYNAPTIC PHARM CORP.
XX
XX Borowsky BE, Ogozalek KL, Jones KA;
XX WPI; 2001-025252/03.
XX
XX Nucleic acid encoding a mammalian (human, rat and mouse) SNORF3 receptor
XX which is useful for designing drugs for treating conditions such as a
XX chronic and acute inflammation, arthritis, neurological disorders and
XX microbial infections.
XX
XX Disclosure; Fig 5; 227p; English.
XX
XX The present invention relates to a mammalian SNORF3 receptor. SNORF3
XX antagonists and agonists are used to treat abnormalities brought about by
XX increased or decreased activity of the mammalian SNORF3 receptor. The
XX receptor is useful as a tool for designing drugs for treating conditions
XX such as a chronic and acute inflammation, arthritis, neurological
XX disorders, microbial infections, bone diseases, respiratory disorders

PI Taupier RJ, Tchernev VT, Vernet CM, Zerhusen BD;
XX WPI; 2002-698672/75.
DR P-PSDB; ADH48916.
XX
PT New NOVX polypeptides or polynucleotides, useful for preventing or
PT treating disorders or syndromes e.g., atherosclerosis, hypertension,
PT obesity or cancer.
XX
PS Claim 8; Page 451; 923pp; English.
XX
CC The present invention relates to novel human NOVX proteins, where X is
CC any number from 1 to 91 and their coding sequences. The proteins and
CC coding sequences are useful for preventing or treating disorders or
CC syndromes e.g. atherosclerosis, hypertension, obesity or cancer. NOV73-
CC NOV91 are GPCR-like proteins. The coding sequence for the present NOV
CC maps to chromosome 6.
XX
SQ Sequence 1039 BP; 283 A; 189 C; 198 G; 369 T; 0 U; 0 Other;

Query Match 99.5%; Score 1032.4; DB 6; Length 1039;
Best Local Similarity 99.9%; Pred. No. 1.9e-271;
Matches 1033; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 5 GAATGATGCCCTTTGGCAAAATTAATTAATTTCTGTGGAAAACAACTGCTCA 64
DB 1 GAATGATGCCCTTTGGCAAAATTAATTAATTTCTGTGGAAAACAACTGCTCA 60
QY 65 ATGATGTCCGTGCTCCCTGTACAGTTTAATGCTGCTCAATTTGACACACCTGTTG 124
DB 61 ATGATGTCCGTGCTCCCTGTACAGTTTAATGCTGCTCAATTTGACACACCTGTTG 120
QY 125 GCAATGCTAGTATTTGTTTCTTATATCACTTCAACAACTTCAACCCCAAAATT 184
DB 121 GCAATGCTAGTATTTGTTTCTTATATCACTTCAACAACTTCAACCCCAAAATT 180
QY 185 GGCTCATTCATTCATGCGCACTGTGCACTTTCTGTGGGGTGTGCTGCTGCTTACA 244
DB 181 GGCTCATTCATTCATGCGCACTGTGCACTTTCTGTGGGGTGTGCTGCTGCTTACA 240
QY 245 GTATGTGATGATCTGCTGAGCACTGTGTTATTTGAGAAAGCTTCTGTAAATTCACA 304
DB 241 GTATGTGATGATCTGCTGAGCACTGTGTTATTTGAGAAAGCTTCTGTAAATTCACA 300
QY 305 CAAGCAGCGACATTAATGCTGAGCTCAGCTCCATTTTCATTTGCTTCATCTCATTTG 364
DB 301 CAAGCAGCGACATTAATGCTGAGCTCAGCTCCATTTTCATTTGCTTCATCTCATTTG 360
QY 365 ACCGCTACTANTGCTGTGTGATCCAGTATTAAGCCAAAGATGAATCTTGTTA 424
DB 361 ACCGCTACTANTGCTGTGTGATCCAGTATTAAGCCAAAGATGAATCTTGTTA 420
QY 425 TTGTGTATGATCTTCAATTAATGAGTGTCCCTGCTGTTTTTGCAATTTGAAATGATCT 484
DB 421 TTGTGTATGATCTTCAATTAATGAGTGTCCCTGCTGTTTTTGCAATTTGAAATGATCT 480
QY 485 TTCTGAGCTAAACTTCAAGGCGCTGAAGATTAATTAACAACATTTTACCTGCAGAG 544
DB 481 TTCTGAGCTAAACTTCAAGGCGCTGAAGATTAATTAACAACATTTTACCTGCAGAG 540
QY 545 GAGGTGCTGTCTTCTTTAGCAAAATATCTGGGGTACTGACCTTTATGACTCTTTT 604
DB 541 GAGGTGCTGTCTTCTTTAGCAAAATATCTGGGGTACTGACCTTTATGACTCTTTT 600
QY 605 ATATACCGGATCTATTAATGTTATGTGTCTATTAACAATTAATCTTATGCTAAAGAAC 664
DB 601 ATATACCGGATCTATTAATGTTATGTGTCTATTAACAATTAATCTTATGCTAAAGAAC 660
QY 665 AGGCAAGATTAATTAAGTATGATCCAAATGCAAGCTCCAAATTTGATGGAATGAAAATG 724
DB 661 AGGCAAGATTAATTAAGTATGATCCAAATGCAAGCTCCAAATTTGATGGAATGAAAATG 720
QY 725 GAATTTCAACAAGCAAGAAAGAAAGCTGTGAAGACATTTGGGATTTGTATGGAGTTT 784
```

```
DB 721 GAATTTCAACAAGCAAGAAAGAAAGCTGTGAAGACATTTGGGATTTGTATGGAGTTT 780
QY 785 TCTATATATGCTGTGCGCCCTTCTTATCTGTACAGTATGAGACCTTTTCTCACTCA 844
DB 781 TCTATATATGCTGTGCGCCCTTCTTATCTGTACAGTATGAGACCTTTTCTCACTCA 840
QY 845 TTATTCACCTTCTTTGAATGATGTGTGATTTGGTGGCTACTTGAACCTCTACATTTA 904
DB 841 TTATTCACCTTCTTTGAATGATGTGTGATTTGGTGGCTACTTGAACCTCTACATTTA 900
QY 905 ATCCAAATGTTATGATTTTCTATCTGTGTTAGAAAAGCACTGAAGATGATGCTGT 964
DB 901 ATCCAAATGTTATGATTTTCTATCTGTGTTAGAAAAGCACTGAAGATGATGCTGT 960
QY 965 TTGTTAAATTTTCCAAAAGATTCATCAGGTGTAATTTATTTTGAATGAGTTGAT 1024
DB 961 TTGTTAAATTTTCCAAAAGATTCATCAGGTGTAATTTATTTTGAATTTGATTTGAT 1020
QY 1025 AGAATTATTAATTT 1038
DB 1021 AGAATTATTAATTT 1034
```

```
RESULT 4
ADH48905
ID ADH48905 standard; DNA; 1039 BP.
XX ADH48905;
XX
XX 25-MAR-2004 (first entry)
XX
XX DE NOV81A coding sequence, SEQ ID 189.
XX
XX KW Human; NOVX; atherosclerosis; hypertension; obesity; cancer; cytostatic;
XX KW hypotensive; antiarteriosclerotic; anorectic; gene therapy; NOV81A;
XX KW GPCR-like protein; chromosome 6; gene; ds.
XX
XX OS Homo sapiens.
XX
XX PN NC0200268652-A2.
XX
XX PD 06-SEP-2002.
XX
XX PF 26-FEB-2002; 2002MO-US005910.
XX
XX PR 26-FEB-2001; 2001US-0271646P.
XX PR 27-FEB-2001; 2001US-0271840P.
XX PR 28-FEB-2001; 2001US-0272404P.
XX PR 28-FEB-2001; 2001US-0272405P.
XX PR 28-FEB-2001; 2001US-0272410P.
XX PR 28-FEB-2001; 2001US-0272414P.
XX PR 02-MAR-2001; 2001US-0272787P.
XX PR 02-MAR-2001; 2001US-0272922P.
XX PR 02-MAR-2001; 2001US-0273048P.
XX PR 02-MAR-2001; 2001US-0273300P.
XX PR 16-MAR-2001; 2001US-0276401P.
XX PR 20-MAR-2001; 2001US-0277324P.
XX PR 20-MAR-2001; 2001US-0278660P.
XX PR 30-MAR-2001; 2001US-0280039P.
XX PR 02-APR-2001; 2001US-0280234P.
XX PR 12-APR-2001; 2001US-0280818P.
XX PR 23-APR-2001; 2001US-0285743P.
XX PR 24-APR-2001; 2001US-0286096P.
XX PR 03-MAY-2001; 2001US-0288353P.
XX PR 17-MAY-2001; 2001US-0291703P.
XX PR 31-MAY-2001; 2001US-0294834P.
XX PR 20-JUN-2001; 2001US-0299695P.
XX PR 21-JUN-2001; 2001US-0299845P.
XX PR 05-JUL-2001; 2001US-0303242P.
XX PR 13-AUG-2001; 2001US-0311981P.
XX PR 16-AUG-2001; 2001US-0312858P.
```


PR 28-FEB-2000; 2000US-0185421P.
 PR 28-FEB-2000; 2000US-0185534P.
 PR 02-MAR-2000; 2000US-0186554P.
 PR 03-MAR-2000; 2000US-0186530P.
 PR 03-MAR-2000; 2000US-0186811P.
 PR 09-MAR-2000; 2000US-0186811P.
 PR 17-MAR-2000; 2000US-0186114P.
 PR 21-MAR-2000; 2000US-0190310P.
 PR 20-APR-2000; 2000US-0190800P.
 PR 02-MAY-2000; 2000US-0198568P.
 PR 08-MAY-2000; 2000US-0201190P.
 PR 25-MAY-2000; 2000US-0203111P.
 XX
 PA (PHMA) PHARMACIA & UPJOHN CO.
 XX
 P1 Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J.
 P1 Schellin KA, Kaytes PS, Bannigan CM, Rufe V, Sejlitz T, Huff RM;
 XX
 DR WPI; 2001-389826/41.
 DR P-PSDB; MAG80973.
 XX
 PT New G protein-coupled receptor (ngPCR-x) and its encoding polynucleotide
 PT useful for diagnosing and treating e.g. schizophrenia.
 XX
 PS Claim 4; Page 92; 261pp; English.

CC The present invention relates to novel G protein-coupled receptors
 CC (ngPCR-x; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27, 28,
 CC 31-38, 40, 41, 53-60) and their coding sequences. The present sequence is
 CC the coding sequence for one such G protein-coupled receptor. GPCRs are
 CC also known as seven transmembrane receptors and function in signal
 CC transduction. The ngPCR coding sequences are useful for screening a
 CC human to diagnose a disorder affecting the brain or a genetic
 CC predisposition, specifically schizophrenia. ngPCR are useful for
 CC identifying compounds useful for treating schizophrenia. Detection of
 CC ngPCR in a sample is useful as a diagnostic tool for diseases or
 CC disorders e.g. thyroid disorders, renal failure, rheumatoid arthritis,
 CC CNS disorders, infectious such as HIV-1, metabolic and cardiovascular
 CC diseases, proliferative disorders and hormonal disorders. Modulators of
 CC ngPCR activity have the utility for treating neurological disorders,
 CC including schizophrenia, ADHD/ADD (attention deficit-hyperactivity
 CC disorder/attention deficit disorder), and neuronal disorders such as
 CC Alzheimer's disease, Parkinson's disease, migraine and senile dementia.
 CC Additional disorders including inflammatory conditions (e.g. Crohn's
 CC disease), rheumatoid arthritis, autoimmune disorders, cancers,
 CC respiratory ailments such as asthma, and inflammatory diseases e.g.
 CC inflammatory bowel disease

SO Sequence 1023 BP; 277 A; 188 C; 199 G; 359 T; 0 U; 0 Other;

Query Match 98.4%; Score 1021.4; DB 4; Length 1023;
 Best Local Similarity 99.9%; Pred. No. 1.9e-266;
 Matches 1022; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGAATGATGCTCTTTGGCAATATTAATATTTCTGTGAAAAACAATGTCGA 63
 DB 1 GGAATGATGCTCTTTGGCAATATTAATATTTCTGTGAAAAACAATGTCGA 60
 QY 64 AATGATGCTGCTCTCTGTCAGTTTAAATGTCATTAATTCGACCACTCGTT 123
 DB 61 AATGATGCTGCTCTCTGTCAGTTTAAATGTCATTAATTCGACCACTCGTT 120
 QY 124 GGCATCTGATATGTTATGTTCTATATCACTTCAAAACAATCAATCCCAACAAT 183
 DB 121 GGCATCTGATATGTTATGTTCTATATCACTTCAAAACAATCAATCCCAACAAT 180
 QY 184 TGGCATCTTCAATTCATGCGCACTGTGACCTTTCTTGCGGGTGTGTCATGCTTAC 243
 DB 181 TGGCATCTTCAATTCATGCGCACTGTGACCTTTCTTGCGGGTGTGTCATGCTTAC 240
 QY 244 AGTATGATGATCTGTCAGACATCTGTGATTTTGAAGAAGCTCTGTAATTAATTCAC 303
 DB 241 AGTATGATGATCTGTCAGACATCTGTGATTTTGAAGAAGCTCTGTAATTAATTCAC 300

QY 304 ACAAGCACCAGATTATGCTGAGCTCAGCCTCCATTTCATTTGCTTCAATCCAT 363
 DB 301 ACAAGCACCAGATTATGCTGAGCTCAGCCTCCATTTCATTTGCTTCAATCCAT 360
 QY 364 GACCGCTACTATGCTGTGTGTATCACTGAGATATTAAGCCAGATGAAATCTTGGTT 423
 DB 361 GACCGCTACTATGCTGTGTGTATCACTGAGATATTAAGCCAGATGAAATCTTGGTT 420
 QY 424 ATTTGATGATCTTCAATTAATGAGAGTGCCTGCTGTTTGGATTTGATGATTC 483
 DB 421 ATTTGATGATCTTCAATTAATGAGAGTGCCTGCTGTTTGGATTTGATGATTC 480
 QY 484 TTTCTGAGCTTAACCTTCAAAAGCGCTGAAGATATATTAACAACATGTCAGAGA 543
 DB 481 TTTCTGAGCTTAACCTTCAAAAGCGCTGAAGATATATTAACAACATGTCAGAGA 540
 QY 544 GGAGGTGCTGCTGCTTTTGAAGAAATATCTGGGGTACTGACCTTTATGACTTTT 603
 DB 541 GGAGGTGCTGCTGCTTTTGAAGAAATATCTGGGGTACTGACCTTTATGACTTTT 600
 QY 604 TATATACCTGATCTATATGTTATGCTCTATTAACAATAATCTTATGCTTAAGAA 663
 DB 601 TATATACCTGATCTATATGTTATGCTCTATTAACAATAATCTTATGCTTAAGAA 660
 QY 664 CAGCAAGATTAATTAATGATGATCCAAATCAGAACTCCAAATTTGAAATGAAAT 723
 DB 661 CAGCAAGATTAATTAATGATGATCCAAATCAGAACTCCAAATTTGAAATGAAAT 720
 QY 724 GGAATTTTCAACAAGAAAGAAAGAGCTGTGAAGCATTTGGGGATTTGATGAGGATT 783
 DB 721 GGAATTTTCAACAAGAAAGAAAGAGCTGTGAAGCATTTGGGGATTTGATGAGGATT 780
 QY 784 TTTCTAATATGCTGTGCTCTTTCTTATCTGATGATGAGACCTTTCTTCTACTAC 843
 DB 781 TTTCTAATATGCTGTGCTCTTTCTTATCTGATGATGAGACCTTTCTTCTACTAC 840
 QY 844 ATTATTCACCTACTTGAATGATGATGATTTGTTGGCTACTTGAATCTTCAATTT 903
 DB 841 ATTATTCACCTACTTGAATGATGATGATTTGTTGGCTACTTGAATCTTCAATTT 900
 QY 904 AATCCAAATGTTATGATTTTCTATCTGCTTTTGAAGAAAGCACTGAAGATGATCTG 963
 DB 901 AATCCAAATGTTATGATTTTCTATCTGCTTTTGAAGAAAGCACTGAAGATGATCTG 960
 QY 964 TTTGGTAAATTTTCCAAAAGATTCATCCAGGTGAAATTAATTTTGAATTTGAGTTCA 1023
 DB 961 TTTGGTAAATTTTCCAAAAGATTCATCCAGGTGAAATTAATTTTGAATTTGAGTTCA 1020
 QY 1024 TAG 1026
 DB 1021 TAG 1023

RESULT 8
 ABS70246
 ID ABS70246 standard; DNA; 1023 BP.
 XX
 AC ABS70246;
 XX
 DT 26-NOV-2002 (first entry)
 XX
 DE DNA encoding human G protein-coupled receptor, ngPCR-56, #2.
 XX
 KW Human; gene; de; G protein-coupled receptor; GPCR; ngPCR; beGPCR;
 KW ng protein coupled receptor; communication; serpentine structure;
 KW seven transmembrane receptor; 7TM; mental disorder; diagnosis;
 KW genetic predisposition; brain; immune response; gene therapy;
 KW anxiety disorder; depression; bipolar disorder; schizophrenia;
 KW Huntington's disease; dyskinesia; manic depression; stroke;
 KW Parkinson's disease; Alzheimer's disease; diabetes; inflammation; wound;
 KW tranquiliser.
 XX
 OS Homo sapiens.

XX WO200264789-A1.
 XX 22-AUG-2002.
 XX 14-FEB-2001; 2001WO-US004641.
 XX 14-FEB-2001; 2001WO-US004641.
 XX (PHARMA) PHARMACIA & UPJOHN CO.
 XX Lind P, Parodi LA, Vogel G, Wood LS;
 XX WPI; 2002-674879/72.
 XX P-PEDB; ABG93791.
 PT New nucleic acids and polypeptides of the ng protein-coupled receptor,
 PT useful for treating or diagnosing a mental disorder or a disorder
 PT affecting the brain, e.g. anxiety disorders, schizophrenia, stroke or
 PT Parkinson's disease.
 XX
 XX Example 1; Page 87; 244pp; English.
 XX The invention discloses an isolated human polypeptide, and encoding
 CC nucleic acid, for a G protein-coupled receptor (GPCR), particularly the
 CC ng protein coupled receptor-14 (ngPCR-14). GPCRs are vital in the
 CC communication between cells and their environment and are characterised
 CC by a serpentine structure that passes through the cell membrane seven
 CC times, hence the reason such receptors are sometimes called seven
 CC transmembrane receptors (7TM). The polynucleotides and polypeptides are
 CC useful for identifying an ngPCR allelic variant that correlates with a
 CC mental disorder, for isolating an antibody that binds to an epitope of
 CC the polypeptide, for identifying a compound that binds the polypeptide or
 CC polynucleotide and/or modulates its biological activity, for screening a
 CC human subject to diagnose a disorder, or a genetic predisposition to a
 CC disorder, affecting the brain or a genetic disposition to the disorder,
 CC for identifying compounds useful for the treatment of a mental disorder
 CC and for identifying a compound useful as a modulator of binding between
 CC ngPCR-14 and a binding partner of ngPCR-14. The polypeptide is also
 CC useful for inducing an immune response in a mammal. The nucleic acid or
 CC polypeptide is particularly useful, using gene therapy, for treating e.g.
 CC anxiety disorders, depression, bipolar disorder, schizophrenia,
 CC Huntington's disease, dyskinesias, manic depression, stroke, Parkinson's
 CC disease or Alzheimer's disease. The nucleic acid and polypeptide may also
 CC be used for treating diabetes, inflammation or wounds. The sequences
 CC presented in ABS70202-ABS70248, ABS70338 and ABS70243 are the DNAs
 CC encoding the ngPCRs (also referred to as beGPCRs)
 CC
 CC Sequence 1023 BP; 277 A; 188 C; 159 G; 359 T; 0 U; 0 Other;
 SQ
 Query Match 98.4%; Score 1021.4; DB 6; Length 1023;
 Best Local Similarity 99.9%; Pred. No. 1.9e-268;
 Matches 1023; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 304 ACAAGCAGCAGCATATGCTGAGCTCAGCCTCCATTTTCCATTGCTTTCATCTCCANT 363
 DB 301 ACAAGCAGCAGCATATGCTGAGCTCAGCCTCCATTTTCCATTGCTTTCATCTCCANT 360
 QY 364 GACCGCTACTATGCTGCTGTGATCCACTGAGATATTAAGCCAGATGATATCTTGTT 423
 DB 361 GACCGCTACTATGCTGCTGTGATCCACTGAGATATTAAGCCAGATGATATCTTGTT 420
 QY 424 ATTTGTGAGATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 483
 DB 421 ATTTGTGAGATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 QY 484 TTTCTGAGCTTAACTTCAAGGCGCTGAAGATATATTAACAATGTTCACTGACAGA 543
 DB 481 TTTCTGAGCTTAACTTCAAGGCGCTGAAGATATATTAACAATGTTCACTGACAGA 540
 QY 544 GGAGGTTGCTGCTGCTTCTTTAGCAAAATATCTGGGGTACGACCTTTATGACTTCTTT 603
 DB 541 GGAGGTTGCTGCTGCTTCTTTAGCAAAATATCTGGGGTACGACCTTTATGACTTCTTT 600
 QY 604 TATATACCTGATCTATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 663
 DB 601 TATATACCTGATCTATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 QY 664 CAGGCAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 723
 DB 661 CAGGCAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 QY 724 GGAATTTGCAAAAGCAAGAAAGAAAGCTGGAAGCACTGGGATGATGATGATGATGAT 783
 DB 721 GGAATTTGCAAAAGCAAGAAAGAAAGCTGGAAGCACTGGGATGATGATGATGATGAT 780
 QY 784 TTCCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 843
 DB 781 TTCCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 QY 844 ATTATTCACCACTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 903
 DB 841 ATTATTCACCACTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
 QY 904 AATCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 963
 DB 901 AATCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 QY 964 TTTGGTAAATTTTCCAAAAGATTCATCCAGGTGTAATTTTGGAAATGAGTTCA 1023
 DB 961 TTTGGTAAATTTTCCAAAAGATTCATCCAGGTGTAATTTTGGAAATGAGTTCA 1020
 QY 1024 TAG 1026
 DB 1021 TAG 1023
 RESULT 9
 ID AB242578 strand; DNA; 1020 BP.
 XX AB242578;
 XX 04-MAR-2003 (first entry)
 DE Human trace amine receptor 1 nucleotide seq ID NO:638.
 XX G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
 KW G protein-coupled receptor modulator; antibody; immune-related disease;
 KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
 KW immunological-related cell proliferative disease; autoimmune disease;
 KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
 KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
 KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
 KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
 KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;

KM hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW ulcer; gene; ds.
XX
XX Homo sapiens.
XX
XX WO200261087-A2.
XX
XX
XX 08-AUG-2002.
XX
XX 19-DEC-2001; 2001WO-US050107.
XX
XX 19-DEC-2000; 2000US-0257144P.
XX
XX (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
XX
XX Burner GC, Roush CL, Brown JP;
XX
XX WPI; 2003-046718/04.
XX
XX P-PSDB; ABP81732.
XX
XX
XX New isolated antigenic peptides e.g., for G protein-coupled receptors
XX (GPCR), useful for diagnosing and designing drugs for treating conditions
XX in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
XX autoimmune diseases.
XX
XX
XX Disclosure; Fig 1; 523bp; English.
XX
XX
XX The present invention describes antigenic peptides (I) comprising: (a)
XX any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
XX acids. Also described: (1) an assay for the detection of a particular G
XX protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
XX and (2) an isolated antibody having high specificity and high affinity or
XX avidity for a particular GPCR. (I) can be used as GPCR modulators and in
XX gene therapy. The antigenic peptides for GPCRs are useful in detecting an
XX antibody against a particular GPCR, and in the production of specific
XX antibodies. The peptides and antibodies are also useful for detecting the
XX presence or absence of corresponding GPCRs. The antigenic peptides for
XX GPCRs and antibodies are useful for diagnosing and designing drugs for
XX treating immune-related diseases, growth-related diseases, cell
XX regeneration-related diseases, immunological-related cell proliferative
XX diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
XX atherosclerosis, bacterial, fungal, protozoan or viral infections,
XX osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
XX inflammation, allergies, Crohn's disease, diabetes, graft versus host
XX disease, Parkinson's disease, multiple sclerosis, psoriasis, host
XX anxiety, depression, schizophrenia, dementia, mental retardation, memory
XX loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
XX hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
XX any other disorder in which GPCRs are involved. The antibodies may be
XX used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
XX GPCR proteins given in ABP81675 to ABP82018, which are used in the
XX exemplification of the present invention
XX
XX
XX Sequence 1020 BP; 275 A; 188 C; 198 G; 359 T; 0 U; 0 Other;
SQ
Query Match 98.3%; Score 1020; DB 8; Length 1020;
Best Local Similarity 100.0%; Pred. No. 4,7e-268;
Matches 1020; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 ATGATGCCCTTTGGCACAATATTAATATTTCTGTGTGAAAAACAACCTGCTCAAT 66
DB 1 ATGATGCCCTTTGGCACAATATTAATATTTCTGTGTGAAAAACAACCTGCTCAAT 60
QY 67 GATGTCGTCGTCCTGTCAGTTTAAATGTCATTAATTTTGACACCACTGTTGGC 126
DB 61 GATGTCGTCGTCCTGTCAGTTTAAATGTCATTAATTTTGACACCACTGTTGGC 120
QY 127 AATGTGATAGTTATGTTTCTATATCAACTTCAAAACAACCTTACCCCAAAATTTG 186
DB 121 AATCTGATAGTTATGTTTCTATATCAACTTCAAAACAACCTTACCCCAAAATTTG 180
QY 187 CTCATTGATTCAGGACCACTGTGACTTTCTTGTGGGGTGTCTGATGACCTTACAGT 246

DB 181 CTCATTGATTCAGGACCACTGTGACTTTCTTGTGGGGTGTCTGATGACCTTACAGT 240
QY 247 ATGTGAGATCTGTGAGCACTGTGGTATTTTGGAGAAGCTTCTGTAATAATACACA 306
DB 241 ATGTGAGATCTGTGAGCACTGTGGTATTTTGGAGAAGCTTCTGTAATAATACACA 300
QY 307 AGCACCGACATTATGCTGAGCTGACCTCCATTTCATTGTGTCTTTCATCTCATTTGAC 366
DB 301 AGCACCGACATTATGCTGAGCTGACCTCCATTTCATTGTGTCTTTCATCTCATTTGAC 360
QY 367 CGCTACTATGCTGTGTGATGACACTGAGATATAAAGCAAGAAATATCTTGGTATT 426
DB 361 CGCTACTATGCTGTGTGATGACACTGAGATATAAAGCAAGAAATATCTTGGTATT 420
QY 427 TGTGTGATGATCTTCATTAGTTGAGTGTCCCTGCTGTTTGGATTGGAATGATCTTT 486
DB 421 TGTGTGATGATCTTCATTAGTTGAGTGTCCCTGCTGTTTGGATTGGAATGATCTTT 480
QY 487 CTGAGAGCTAACTTCAAGAGGCGCTGGAAGATATATTACAAACATGTTCACTGAGAGA 546
DB 481 CTGAGAGCTAACTTCAAGAGGCGCTGGAAGATATATTACAAACATGTTCACTGAGAGA 540
QY 547 GGTGCTCTGCTCTTCTTAGCAAAATATCTGGGTAAGTACCTTATGACTCTTTTAT 606
DB 541 GGTGCTCTGCTCTTCTTAGCAAAATATCTGGGTAAGTACCTTATGACTCTTTTAT 600
QY 607 ATACTGATCTATTATGTTATGTGTCTATTACAGAAATATATCTTATCGTTAAAGACAG 666
DB 601 ATACTGATCTATTATGTTATGTGTCTATTACAGAAATATATCTTATCGTTAAAGACAG 660
QY 667 GCAGATTAATTTAGTGTGATGCAATCAGACCTCCAAATTTGGAATGGAATGGAATGGA 726
DB 661 GCAGATTAATTTAGTGTGATGCAATCAGACCTCCAAATTTGGAATGGAATGGAATGGA 720
QY 727 ATTTCACAAAGCAAGAAAGAAAGCTGTGAAGCATTTGGGATTTGATGGAGTTTTC 786
DB 721 ATTTCACAAAGCAAGAAAGAAAGCTGTGAAGCATTTGGGATTTGATGGAGTTTTC 780
QY 787 CTATATGCTGTGTCCTTTCTTATCTGTACAGTATGACCTTTTCTTCACTACAT 846
DB 781 CTATATGCTGTGTCCTTTCTTATCTGTACAGTATGACCTTTTCTTCACTACAT 840
QY 847 ATTCCACCTTATGATGATGATGTTGATTTGGCTTGAACCTTACATTAAT 906
DB 841 ATTCCACCTTATGATGATGATGTTGATTTGGCTTGAACCTTACATTAAT 900
QY 907 CCAATGTTTATGATGATTTTCTATCTTGGTTAGAAAAGCACTGAAGATGATGCTGTT 966
DB 901 CCAATGTTTATGATGATTTTCTATCTTGGTTAGAAAAGCACTGAAGATGATGCTGTT 960
QY 967 GGTAAATTTTCCAAAAGATTCATCCAGGTGAATTTTGGAAATTTGAGTTGATCATAG 1026
DB 961 GGTAAATTTTCCAAAAGATTCATCCAGGTGAATTTTGGAAATTTGAGTTGATCATAG 1020
RESULT 10
ADS73690
ID ADS73690 standard; DNA; 1020 BP.
XX
XX ADS73690;
XX
XX 16-DEC-2004 (first entry)
XX
XX Human trace amine receptor 1 (TA1) polypeptide encoding DNA.
XX
XX
XX TA1; GPCR; trace amine receptor 1; antianaemic; antiinflammatory;
XX cardiovascular; neuroprotective; respiratory; uropathic;
XX G protein coupled receptor; gene therapy; human; CSRL; gene; ds.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX FH 1..1020
XX CDS


```
/*tag= a
/product= "TA1 receptor"
/note= "trace amine receptor 1"
```

WO2004083851-A2.

30-SEP-2004.

04-MAR-2004; 2004WO-EP002182.

18-MAR-2003; 2003EP-00006101.

(FARB) BAYER HEALTHCARE AG.

Golz S, Brueggemeier U, Geerts A;

WPI; 2004-691179/67.

GENBANK; NM_138327.

Screening for thera

Screening for therapeutic agents, useful for treating e.g., cardiovascular and inflammatory disorders, comprises contacting a test compound with trace amine receptor 1 (TA1) polypeptide and detecting their binding.

Disclosure; SEQ ID NO 1; 117pp; English.

The invention relates to screening for therapeutic agents and involves contacting a test compound with trace amine receptor 1 (TAR) polypeptide or polynucleotide and detecting binding of the test compound to TAR polypeptide or polynucleotide, or determining TAR polypeptide activity at a certain test compound concentration or in the absence of the test compound and at a different concentration of the test compound. The therapeutic agents are useful in treating disease such as cardiovascular disorders, inflammatory diseases, hematological disorders, respiratory diseases, neurological disorders or urological disorders in a mammal. The regulators of TAR are useful in the preparation of a pharmaceutical composition for treating the above diseases and also useful for the regulation of TAR activity in a mammal having the disease. The nucleotide sequences encoding TAR are useful as hybridization probes, in constructing oligomers for PCR, for chromosome and gene mapping, in the recombinant production of TAR, in generating antisense DNA or RNA and in molecular biology techniques that have not yet been developed. TAR are useful for immunizing a mammal to produce polyclonal antibodies and for diagnostic purposes. The present sequence represents a DNA encoding a human TAR receptor, a seven transmembrane G protein coupled receptor polypeptide.

Sequence 1020 BP; 275 A; 188 C; 198 G; 359 T; 0 U; 0 Other;

Query Match 1020;	98.3%;	Score 1020;	DB 13;	Length 1020;
Best Local Similarity	100.0%;	Pred. No. 4.7e-268;		
Matches 1020;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

Qy	7	ATGATGCCCTTTGGCACAATATTAATTAATTTCTGTGTGAAAAACAACCTGCTCAAT	66
Db	1	ATGATGCCCTTTGGCACAATATTAATTAATTTCTGTGTGAAAAACAACCTGCTCAAT	60
Qy	67	GATGTCCGCTCTCCCTGTACAGTTTAATGTGTCTATATTTCTGACCACTCTGTTGGC	126
Db	61	GATGTCCGCTCTCCCTGTACAGTTTAATGTGTCTATATTTCTGACCACTCTGTTGGC	120
Qy	127	AATCGATAGTAATTTGTTCTATATACACTTCAACAACCTCATATCCCCAATAATTGG	186
Db	121	AATCGATAGTAATTTGTTCTATATACACTTCAACAACCTCATATCCCCAATAATTGG	180
Qy	187	CTCATTCATTCATGGCCCACTGTGCACTTCTTCTGTGGGTGTCTGTATCGCTTACAGT	246
Db	181	CTCATTCATTCATGGCCCACTGTGCACTTCTTCTGTGGGTGTCTGTATCGCTTACAGT	240
Qy	247	ATGTGTAGATCTGTCTGAGCACTGTGTGTAATTTTGGAGAAAGTCTTCTGTATAATTACACA	306
Db	241	ATGTGTAGATCTGTCTGAGCACTGTGTGTAATTTTGGAGAAAGTCTTCTGTATAATTACACA	300

QY	30	AGACAGCAACATTAGCTGAGCTGAGCGCTCCATTTCGATTGCTTCATCTCCATTGAC	366
Db	301	AGCACCGACATTATGCTGAGCTGAGCGCTCCATTTCGATTGCTTCATCTCCATTGAC	360
QY	367	CGCTACTATGCTGTGTGTGATCCACTGAGATATATAAGCCAAATGAATATCTTGTTATT	426
Db	361	CGCTACTATGCTGTGTGTGATCCACTGAGATATATAAGCCAAATGAATATCTTGTTATT	420
QY	427	TGTGTGATGATCTTCATTAGTTGGAGTGTCCCTGCTGTGTTTTTGATTTGGAAATGATCTTT	486
Db	421	TGTGTGATGATCTTCATTAGTTGGAGTGTCCCTGCTGTGTTTTTGATTTGGAAATGATCTTT	480
QY	487	CTGGAGCTAAACCTTCAAAGGCGCTGAGAGAGATATATTACAAACATGTTCACTGACAGAGA	546
Db	481	CTGGAGCTAAACCTTCAAAGGCGCTGAGAGAGATATATTACAAACATGTTCACTGACAGAGA	540
QY	547	GGTGTGCTGTGCTCTCTTTAGCAAAATATCTGGGGTACTGACCTTTATGACTCTCTTTTAT	606
Db	541	GGTGTGCTGTGCTCTCTTTAGCAAAATATCTGGGGTACTGACCTTTATGACTCTCTTTTAT	600
QY	607	ATTACTGATCTATTATGTTATGTGTCTATTATACGAATATATCTTATCCGTTAAAGAACAG	666
Db	601	ATTACTGATCTATTATGTTATGTGTCTATTATACGAATATATCTTATCCGTTAAAGAACAG	660
QY	667	GCAAGATTAAATTAGTANGCCAAATCAGAGCTCCAAATTGGAATTGAAATGAAATAATGGA	726
Db	661	GCAAGATTAAATTAGTANGCCAAATCAGAGCTCCAAATTGGAATTGAAATGAAATAATGGA	720
QY	727	ATTTCACAAAGCAAAAGAAAGAAAGCTGTGAGACATTTGGGGATTTGTATGGAGATTTC	786
Db	721	ATTTCACAAAGCAAAAGAAAGAAAGCTGTGAGACATTTGGGGATTTGTATGGAGATTTC	780
QY	787	CTATATGCTGTGTCGCCCTTTCTTTATCTGTACAGTACAGACCTTTTCTTCACTACATT	846
Db	781	CTATATGCTGTGTCGCCCTTTCTTTATCTGTACAGTACAGACCTTTTCTTCACTACATT	840
QY	847	ATTCCACCTACTTTGANGANGTGTGATTGGTTGGTCTACTTGAACCTACATTTAAT	906
Db	841	ATTCCACCTACTTTGANGANGTGTGATTGGTTGGTCTACTTGAACCTACATTTAAT	900
QY	907	CCAATGGTTATATGCAATTTTCTATCTTGTTGTAAGAAAGCAGTGAAGATGATGCTGTT	966
Db	901	CCAATGGTTATATGCAATTTTCTATCTTGTTGTAAGAAAGCAGTGAAGATGATGCTGTT	960
QY	967	GGTAAATTTTCCAAAAGATTCATCCAGGTGTAAATTTATTTTGGAAATGAGTTCAATAG	1026
Db	961	GGTAAATTTTCCAAAAGATTCATCCAGGTGTAAATTTATTTTGGAAATGAGTTCAATAG	1020
RESULT 11			
ABT04871			
ID	ABT04871	standard; cDNA; 1020 BP.	
XX	AC	ABT04871;	
XX	DT	11-OCT-2002 (first entry)	
XX	DE	Human G protein coupled receptor hRUP3 coding sequence.	
XX	KW	Human; G-protein coupled receptor; GPCR; hRUP28; hRUP29; hRUP30; hRUP31;	
XX	KX	hRUP32; hRUP33; hRUP34; hRUP35; hRUP36; hRUP37; gene; ss.	
XX	OS	Homo sapiens.	
XX	PN	WO200242461-A2.	
XX	PD	30-MAY-2002.	
XX	PE	26-NOV-2001; 2001WO-US044386.	
XX	27-NOV-2000;	2000US-0253404P.	
PR	12-DEC-2000;	2000US-0255366P.	

DR P-PsDB; ADO29662.
XX Novel mammalian G protein coupled receptors, useful for identifying
PT compounds that mediate disease diagnosis and treating disease condition
PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
PT pectoris, Parkinson's disease.
XX
XX Claim 151; SEQ ID NO 1152; 542bp; English.
XX
CC The invention relates to human and mouse G protein-coupled receptors
CC (GPCRs) and nucleic acids encoding them. The invention also relates to
CC sequences at least 90% identical to the GPCR proteins and nucleic acids
CC of the invention; methods of treating, preventing or diagnosing diseases
CC associated with GPCRs of the invention; methods of screening for
CC compounds useful in the treatment of GPCR-related diseases; a transgenic
CC mouse comprising a GPCR gene of the invention; a mouse comprising a
CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
CC from the transgenic mice; kits comprising several mice, each of which has
CC a mutation in a different GPCR gene of the invention; and kits comprising
CC probes which hybridize to GPCR polynucleotides of the invention. The
CC invention further discloses variants of the GPCR polypeptides and vectors
CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
CC be used in the diagnosis, treatment or prevention of a wide variety of
CC diseases including neurological disorders (e.g., Alzheimer's disease,
CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
CC disorders of the adrenal gland; disorders of the colon or intestine
CC (e.g., Crohn's disease, diarrhea, food poisoning or irritable bowel
CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
CC myocardial infarction); muscular disorders; blood disorders (e.g.,
CC anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
CC thyroid (e.g., cancers). The present sequence represents a GPCR-encoding
CC nucleic acid of the invention. Note: The full sequence data for this
CC patent did not form part of the printed specification; those sequences
CC not shown were obtained in electronic format directly from WIP0 at
CC ftp.wipo.int/pub/published_pcr_sequences.
XX
XX Sequence 1020 BP; 276 A; 188 C; 197 G; 359 T; 0 U; 0 Other;
XX
Query Match 98.1%; Score 1018.4; DB 12; Length 1020;
Best Local Similarity 99.9%; Pred. No. 1.3e-267;
Matches 1019; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 7 ATGATGCCCTTTGGCACAATATTAATTAATTTCTGTGTGAAAAACAACGTGTCAAAT 66
DB 1 ATGATGCCCTTTGGCACAATATTAATTAATTTCTGTGTGAAAAACAACGTGTCAAAT 60
OY 67 GATGCCGTGCTTCCCTGTACAGTTTATGTCATATTTGTGACCACTGCTGGC 126
DB 61 GATGCCGTGCTTCCCTGTACAGTTTATGTCATATTTGTGACCACTGCTGGC 120
OY 127 AATCGATAGTATGTTTCTATACACTTCAACAACCTCATACCCCAACAATTGG 186
DB 121 AATCGATAGTATGTTTCTATATCACTTCAACAACCTCATACCCCAACAATTGG 180
OY 187 CTCATTCAATCCATGGCCACTGTGACCTTTCTTGGGGTGTCTGTCATGCTTACAGT 246
DB 181 CTCATTCAATCCATGGCCACTGTGACCTTTCTTGGGGTGTCTGTCATGCTTACAGT 240
OY 247 ATGTCGAAGTCTGTCGAGACACTGTTGTTATTTTGGAGAACTTCTGTTAAATTACACA 306
DB 241 ATGTCGAAGTCTGTCGAGACACTGTTGTTATTTTGGAGAACTTCTGTTAAATTACACA 300
OY 307 AGCAGCGACATTATGTCGAGCTCAGCTGCATTTTCCATTTGCTTTCATCCATTGAC 366
DB 301 AGCAGCGACATTATGTCGAGCTCAGCTGCATTTTCCATTTGCTTTCATCCATTGAC 360
OY 367 CGCTACTATGCTGTGTGTGATCCACTGAGATTAAGCCAGATGAATATCTTGTTAT 426

DB 361 CGTACTATGCTGTGTGTGATCCACTGAGATATAAAGCCAAAGATGATATCTTGTTAT 420
OY 427 TGTGTGATGATCTTCATTAAGTTGAGAGTCCCTGCTGTTTTCATTTGGAATGATCTT 486
DB 421 TGTGTGATGATCTTCATTAAGTTGAGAGTCCCTGCTGTTTTCATTTGGAATGATCTT 480
OY 487 CTGAGACTAACTTCAAGAGCGCTGAGAGATATATTAACAACATGTTCACTGACAGAGA 546
DB 481 CTGAGACTAACTTCAAGAGCGCTGAGAGATATATTAACAACATGTTCACTGACAGAGA 540
OY 547 GATTCCTGTCTTCTTTAGCAAAATATCTGGGGTACTGACCTTTATGATCTTTTAT 606
DB 541 GATTCCTGTCTTCTTTAGCAAAATATCTGGGGTACTGACCTTTATGATCTTTTAT 600
OY 607 ATACTGTGATCTATTAATGTTATGTTCTATTAAGATATATCTATGCTTAAGAACAG 666
DB 601 ATACTGTGATCTATTAATGTTATGTTCTATTAAGATATATCTATGCTTAAGAACAG 660
OY 667 GCAAGATTAATTAAGATGACCAATCAGAAAGCTCCAAATTTGATGGAATGAAATGGA 726
DB 661 GCAAGATTAATTAAGATGACCAATCAGAAAGCTCCAAATTTGATGGAATGAAATGGA 720
OY 727 ATTTCAAGACAAAGAAAGAAAGAGCTGTGAAGACATTTGGGATTTGTATGAGAGTTTC 786
DB 721 ATTTCAAGACAAAGAAAGAAAGAGCTGTGAAGACATTTGGGATTTGTATGAGAGTTTC 780
OY 787 CTAATATGCTGTGCTTTCTTTATCTGTACAGTACAGACCTTTCTTACATCAAT 846
DB 781 CTAATATGCTGTGCTTTCTTTATCTGTACAGTACAGACCTTTCTTACATCAAT 840
OY 847 ATTCACCTACTTGAATGATGATGTGTTGTTGTTGCTACTTGAACCTTACATTTAT 906
DB 841 ATTCACCTACTTGAATGATGATGTGTTGTTGTTGCTACTTGAACCTTACATTTAT 900
OY 907 CCAATGTTTATGCAATTTTCTATCTTGTGTTAGAAAAGCACTGAGATGCTGTTT 966
DB 901 CCAATGTTTATGCAATTTTCTATCTTGTGTTAGAAAAGCACTGAGATGCTGTTT 960
OY 967 GGTAAATTTTCCAAAAGATTCATCCAGGTGAATATTTTGGAAATGATGCTCATAG 1026
DB 961 GGTAAATTTTCCAAAAGATTCATCCAGGTGAATATTTTGGAAATGATGCTCATAG 1020

RESULT 13
AAH26919
ID AAH26919 standard; cDNA, 1017 BP.
XX
XX AAH26919;
XX
XX 21-DEC-2001 (first entry)
XX
XX Human G protein coupled receptor cDNA.
XX
XX G protein coupled receptor; GPCR; aminergic receptor;
XX signal transduction; gene therapy; human; ss.
XX
XX Homo sapiens.
XX
XX WO200172841-A2.
XX
XX 04-OCT-2001.
XX
XX 27-MAR-2001; 2001WO-US009660.
XX
XX 27-MAR-2000; 2000US-0192311P.
XX PR 04-AUG-2000; 2000US-00633145.
XX
XX (PEKE) PE CORP NY.
XX
XX Kodira C, Cravchik A, Di Francesco V, Beasley EM;
XX WPI, 2001-611617/70.
XX
XX P-PsDB; AAB82970.
DR

PS Example 1; Page 86; 244pp; English.
XY

the invention discloses an isolated human polypeptide, and encoding nucleic acid, for a G protein-coupled human receptor (GPCR), particularly the ng protein coupled receptor-14 (NGPCR-14). GPCRs are vital in the communication between cells and their environment and are characterized by a serpentine structure that passes through the cell membrane seven times, hence the reason such receptors are sometimes called seven transmembrane receptors (7TM). The polynucleotides and polypeptides are useful for identifying an ngPCR allelic variant that correlates with a mental disorder, for isolating an antibody that binds to an epitope of the polypeptide, for identifying a compound that binds the polypeptide or polynucleotide and/or modulates its biological activity, for screening a human subject to diagnose a disorder, or a genetic predisposition to a disorder, affecting the brain or a genetic disposition to the disorder, for identifying compounds useful for the treatment of a mental disorder and for identifying a compound useful as a modulator of binding between ngPCR-14 and a binding partner of ngPCR-14. The polypeptide is also useful for inducing an immune response in a mammal. The nucleic acid or polypeptide is particularly useful, using gene therapy, for treating e.g., anxiety disorders, depression, bipolar disorder, schizophrenia, Huntington's disease, dyskinesias, manic depression, stroke, Parkinson's disease or Alzheimer's disease. The nucleic acid and polypeptide may also be used for treating diabetes, inflammation or wounds. The sequences presented in *AB570202-AB570248*, *AB570338* and *AB570243* are the DNAs encoding the ngPCRs (also referred to as *ngPCRs*)

Sequence 1138 BP; 339 A; 207 C; 220 G; 372 T; 0 U; 0 Other;

Query Match	Score	DB	Length
83.7%;	869;	6;	1138;
Best Local Similarity	100.0%	1	1138

Matches 869; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TCAGAAATGATGACCCCTTTGSCAACATATATATTAATTTCTGTGTGTAAGAAAACA	CTGG	60
Db	269	TCAGAAATGATGCCCTTTTGCACAAATATATTAATTAATTTCTGTGTGTAAGAAAACA	CTGG	328
QY	61	TCAAATGATGTCGGTGCTTCCGTGACAGTTTAATGTCCTCATTAATCTGACCA	CTC	128
Db	329	TCAAATGATGTCGGTGCTTCCGTGACAGTTTAATGTCCTCATTAATCTGACCA	CTC	388
QY	121	GTTGGCAATCTGATAGTATTTGTTTCTATATCACTTCAAACAATTCATACCCACA	180	
Db	389	GTTGGCAATCTGATAGTATTTGTTTCTATATCACTTCAAACAATTCATACCCACA	448	
QY	181	AATTGCTCATTCATTCACATGAGCACACTGAGACCTTCTTGCGGGTGCTGATGCTT	240	
Db	449	AATTGCTCATTCATTCACATGAGCACACTGAGACCTTCTTGCGGGTGCTGATGCTT	508	
QY	241	TACAGTATGATGAGATCTGCTGAGCACTGTGGTATTTTGGAGAACTTCTGTAAATT	300	
Db	509	TACAGTATGATGAGATCTGCTGAGCACTGTGTGTAATTTTGGAGAACTTCTGTAAATT	568	
QY	301	CACACAAGCACCGACATTAATGCTGAGCTGAGCTCCATTTTCCATTGTCCTTCACTCC	360	
Db	569	CACACAAGCACCGACATTAATGCTGAGCTGAGCTCCATTTTCCATTGTCCTTCACTCC	628	
QY	361	ATTGACCGCTACATATCTGTGTGTATCCACAGATATTAAGCAAGATGAATATCTTG	420	
Db	629	ATTGACCGCTACATATCTGTGTGTATCCACAGATATTAAGCAAGATGAATATCTTG	688	
QY	421	GTTATTTGTGTGATGATCTTCAATAGTTGAGTGCCCTGCTGTTTGGACATTTGGAAATG	480	
Db	689	GTTATTTGTGTGATGATCTTCAATAGTTGAGTGCCCTGCTGTTTGGACATTTGGAAATG	748	
QY	481	ATCTTCTGAGAGCTAACTTCAAAGCGCTGAGAGATATTTACAACAATGTTCACTGC	540	
Db	749	ATCTTCTGAGAGCTAACTTCAAAGCGCTGAGAGATATTTACAACAATGTTCACTGC	808	
QY	541	AGAGGAGTTGCTGTCTCTCTTTTGAAGAAATATCTGGGGGATACGACCTTATGACTTCT	600	
Db	809	AGAGGAGTTGCTGTCTCTCTTTTGAAGAAATATCTGGGGGATACGACCTTATGACTTCT	868	

Qy	601	TTTATATACCTGATCTAATATGTATAGTGTCTATTACAGAAATATATCTTAAGCTTAA	660
Dp	869	TTTATATACCTGATCTAATATGTATAGTGTCTATTACAGAAATATATCTTAAGCTTAA	928
Qy	661	GAACAGGCAAAATTTATTTAGTATGCCAATAGAAAGCTCCAAATTTGATGGAAATGAAA	720
Dp	929	GAACAGGCAAAATTTATTTAGTATGCCAATAGAAAGCTCCAAATTTGATGGAAATGAAA	988
Qy	721	AATGGAATTTCACAAAGCAAAAGAAAGAAAGAAAGCTGTGAAGACATTGGGGAATTGTATGGGA	780
Dp	989	AATGGAATTTCACAAAGCAAAAGAAAGAAAGAAAGCTGTGAAGACATTGGGGAATTGTATGGGA	1048
Qy	781	GTTTTCCTAATATGCTGTGCCCTTTCTTTATCTGTACAGTCATGAGACCCCTTTTCTTGAC	840
Dp	1049	GTTTTCCTAATATGCTGTGCCCTTTCTTTATCTGTACAGTCATGAGACCCCTTTTCTTGAC	1108
Qy	841	TACATTTATTCACCTACTTGTGAATGATGT	869
Dp	1109	TACATTTATTCACCTACTTGTGAATGATGT	1137

Search completed: February 12, 2005, 00:52:33
Job time : 468 secs

Search completed: February 12, 2005, 00:52:33
Job time : 468 secs

QY	154	CAC TTC A A A C A A C T T C A T A T A C C C A A C A A A T T G G C T A T T C A T T C A T T G C A C T G T G G A C	213
Db	61	C A C T T C A A A C A A C T T C A T A T A C C C A A C A A A T T G G C T A T T C A T T C A T T G C C A C T G T G G A C	120
QY	214	T T T C T T C G G G G T G T C T G G T C A T G C C T T A C A G T A T T G G T G A T T G T G T G A C A C T G T T T G G	273
Db	121	T T T C T T C G G G G T G T C T G G T C A T G C C T T A C A G T A T T G G T G A T T G C T G T G A C A C T G T T T G G	180
QY	274	T A T T T T G A G A A G T C T T T C T G T A A A T T C A C A G A C A C C A C A T T A T G C T G A G C T C A G C C	333
Db	181	T A T T T T G A G A A G T C T T G T A A A T T C A C A G A C A C C A C A T T A T G C T G A G C T C A G C C	240
QY	334	T C C A T T T T C C A T T T G T C T T T C A T C T C C A T T G A C C G C T A C T A T G C T G T G T G A T C C A C T G	393
Db	241	T C C A T T T T C C A T T T G T C T T T C A T C C A T T G A C C G C T A C T A T G C T G T G T G A T C C A C T G	300
QY	394	A G A T A T A A A G C C A G A T G A A T A T C T T G T A T T T G T G T G A T G A T C T T C A T T A G T G A G T	453
Db	301	A G A T A T A A A G C C A G A T G A A T A T C T T G T A T T T G T G T G A T G A T C T T C A T T A G T G A G T	360
QY	454	G T C C C G C T G T T T T T G C A T T T G A T G A T G A T C T T T C T G A G C T A A C T T C A A A G C G C T G A A	513
Db	361	G T C C C G C T G T A N	420
QY	514	G A G A T A T A T A C A A C A T G T T C A C T G C A G A G A G G T T G C T G T C T T C T T A G C A A A T A	573
Db	421	N N	480
QY	574	T C T G G G G T A C T G A C C T T T A T G A C T T C T T T T A T A T A C T G G A T C T A T T A T G T A T G T G C	633
Db	481	T C T G G G G T A C T G A C C T T T A T G A C T T C T T T T A T A T A C T G G A T C T A T T A T G T A T G T G C	540
QY	634	T A T T A C A G A T A T C T T A T C G T T A A G A A C G G A A G T T A A T T A T G T G A T G C C A T C A G	693
Db	541	T A T T A C A G A T A T A T C T T A T C G T T A A G A A C G G A A G T T A A T T A T G T A T G C C A T C A G	600
QY	694	A A G C T C C A A T T G A T G A T T G A A A T G A A A A T G A A T T T C A A A G C A A A A A A A G A A A G C T	753
Db	601	A A G C T C C A A T T G A T G A T T G A A A T G A A A A T T T C A A A G C A A A A A A A A A A A A G T	660
QY	754	G T G A A G A C A T T G G G G A T T G T G A T T G G A A G T T T C T A T A T A T G C T G T G C C T T T C T T A T C	813
Db	661	G T G A A G A C A T T G G G G A T T G T G A T T G G A A G T T T C T A T A T A T G C T G T G C C T T T C T T A T C	720
QY	814	T G T A C A G T A C T G A C C T T T T C T T A C T A C A T T A T T T C C A C T A C T T G A A T G A T G T A T G	873
Db	721	T G T A C A G T A C T G A C C T T T T C T T A C T A C A T T A T T T C C A C C T A C T T T G A A T G A T G A T T G	780
QY	874	A T T T G T T G G T A C T T G A A C T C T A C A T T T A T C C A A T G T T T A T A T G A C A T T T T C T A T C C T	933
Db	781	A T T T G T T G G T A C T T G A A C T C T A C A T T T A T C C A A T G T T T A T A T G A C A T T T T T C T A T C C T	840
QY	934	T G G T T T A G A A A A G C A C T A A G A T G A T G C T T T G T T A A A T T T T C A A A A A A A T T C A C C	993
Db	841	T G G T T T A G A A A A G C A C T A A G A T A T G C T T T G T T A A A T T T T C C A A A A A A A T T C A C C	900
QY	994	A G G T G T A A T A T T T T T G A A T T G A T T G A T T G C A T A G	1026
Db	901	A G G T G T A A T A T T T T T T G A A T T G A T T G A T T G C A T A G	933

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	(bases 1 to 881)		
	NIH-MGC	http://mgc.nci.nih.gov/		
	National Institutes of Health, Mammalian Gene Collection (MGC)			
	Unpublished (1999)			
	Contact: Daniela S. Gerhard, Ph.D.			
	Office of Cancer Genomics			
	National Cancer Institute / NIH			
	Bldg. 31 Rm10A07 Bethesda, MD 20892			
	Email: ggsdpd-femail.nih.gov			
	Tissue Procurement: GPCR Consortium			
	cDNA Library Preparation: GPCR Consortium			
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)			
	DNA Sequencing by: Agencourt Bioscience Corporation			
	Clone distribution: MGC clone distribution information can be			
	found through the I.M.A.G.E. Consortium/LNL at:			
	http://image.llnl.gov			
	Plate: IRB14 row: f column: 06			
	High quality sequence stop: 682.			
	Location/Qualifiers			
	1..881			
	/organism="Homo sapiens"			
	/mol_type="mrna"			
	/db_xref="taxon:9606"			
	/clone="IMAGE:30706625"			
	/issue_type="mixed"			
	/lab_host="DH10B"			
	/clone_lib="NIH MGC 145"			
	/note="Vector: pCDNA3.1; Site 1 varies by clone; Site 2 varies by clone; ORFs were PCR-amplified and cloned into pCDNA3.1 by the GPCR Consortium. Cloning sites vary by clone and include the following: 5'-EcoRV-XmiI/XhoI-3', 5'-EcoRV-XmiI/XhoI-3', EcoRV (TA cloned, non-directional)			
	For information about which gene each clones represents, please visit our anonymous ftp site at			
	ftp://image.llnl.gov/image/featraved_plates/IRB1.presv.dat			
	a Note: this is a NIH-MGC Library."			
ORIGIN				
Query Match	67.3%; Score 698.6; DB 7; Length 881;			
Best Local Similarity	90.3%; Pred. NO. 3.8e-167;			
Matches 757; Conservative	0; Mismatches 80; Indels 1; Gaps 1;			
QY	1	TCAGAAATGATGCCCTTTTGCACAAATTAATTAATTTCTGTTGAAAAACAACCTGG	60	
DB	14	TCAGAAATGATGCCCTTTTGCACAAATTAATTAATTTCTGTTGAAAAACAACCTGG	73	
QY	61	TCAAATGATGTCGGGCTTCCCTGTAAGAGTTTAAAGGCTCATTAATTCGACCACTC	120	
DB	74	TCAAATGATGTCGGGCTTCCCTGTAAGAGTTTAAAGGCTCATTAATTCGACCACTC	133	
QY	121	GTGGCAATCTGATGTTATTTGTTCTAATATCACACTTCAACAACCTTCATACCCACA	180	
DB	134	GTGGCAATCTGATGTTATTTGTTCTAATATCACACTTCAACAACCTTCATACCCACA	193	
QY	181	AATGGCTCATTAATTCAGAGCCACGTCGACCTTTCTTCTGAGGCTCTGCTATGCTT	240	
DB	194	AATGGCTCATTAATTCAGAGCCACGTCGACCTTTCTTCTGAGGCTCTGCTATGCTT	253	
QY	241	TACAGTATGATGATCTGCTGAGACACTGTTGTAATTTGAGAAAGCTTCTGTAATTT	300	
DB	254	TACAGTATGATGATCTGCTGAGACACTGTTGTAATTTGAGAAAGCTTCTGTAATTT	313	
QY	301	CACGACGACCCGACATTAATGCTGAGCTCAGCTCCATTTTCCATTTGCTTTATCTCC	360	
DB	314	CACGACGACCCGACATTAATGCTGAGCTCAGCTCCATTTTCCATTTGCTTTATCTCC	373	
QY	361	ATTGACCGCTACTATGCTGTTGATGATCCATGATATTAAGCAAGATATATCTTGG	420	
DB	374	ATTGACCGCTACTATGCTGTTGATGATCCATGATATTAAGCAAGATATATCTTGG	433	
QY	421	GTTATTTTGTGATGATCTTCATTAAGTTGAGAGTGTCCCTGCTTTTGGCATTTGGAAG	480	

Db 434 GTTATTTGTGATGATCTTCAATAGTAGAGTGTCCCTGCTGTTTTTGCAATTTGGAATG 493
QY 481 ATCTTTCTGAGCTAACTTCAAGGCGCTGAAGAGATATATATACAAACATGTCACCTG 540
Db 494 ATCTTTCTGAGCTAACTTCAAGGCGCTGAAGAGATATATATACAAACATGTCACCTG 553
QY 541 AGAGAGAGTGTGCTGCTCTCTTTAGCAAAATATCTGGGTACTGACCTTTATGACTTCT 600
Db 554 AGAGAGAGTGTGCTGCTCTCTTTAGCAAAATATCTGGGTACTGACCTTTATGACTTCT 613
QY 601 TTTTATATACCTGATCTATTTATGTTATGTCCTATTAAGATATATTTATTCCTTAA 660
Db 614 TTTTATATACCTGATCTATTTATGTTATGTCCTATTAAGATATATTTATTCCTTAA 673
QY 661 GAACAGGCAAGATTAATAGTAGTGC--AATCAGAACTCCAAATTTGATGAAATGA 719
Db 674 GAACAGGCAAGATTAATAGTAGTGC--AATCAGAACTCCAAATTTGATGAAATGA 733
QY 720 AAATGAAATTTCAAGCAAGAAAGAAAGAGCTGTGAAGACATTTGGGATTTGATGG 779
Db 734 AAATGAAATTTCAAGCAAGAAAGAAAGAGCTGTGAAGACATTTGGGATTTGATGG 793
QY 780 AGTTTCTCAATATAGTGTGCTGCTTTCTTTATCTGTACAGTCAAGACCTTTTCTT 837
Db 794 ATTTGTTGTATGGGCAAGATTTTCCCTAAATATGCTGGGATGGCCCTTTTCTT 851

RESULT 3

CN832845 824 bp mRNA linear EST 02-JUN-2004

LOCUS CN832845 824 bp mRNA linear EST 02-JUN-2004

DEFINITION AGENCOURT_15669803 NIH_MGC_145 Homo sapiens cDNA clone

IMAGE:30706626 5', mRNA sequence.

ACCESSION CN832845

VERSION CN832845.1 GI:47936624

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 824)

NIH-MGC http://mgs.mci.nih.gov/

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgsapbs-r@mail.nih.gov

Tissue Procurement: GPCR Consortium

cDNA Library Preparation: GPCR Consortium

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: IRB14 row: f column: 07

High quality sequence step: 519.

Location/Qualifiers

1..824

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30706626"

/tissue_type="mixed"

/lab_host="DH10B"

/clone_lib="NIH_MGC_145"

/note="Vector: pCDNA3.1; Site 1: varies by clone; Site 2:

varies by clone; ORFs were PCR-amplified and cloned into

pCDNA3.1 by the GPCR Consortium. Cloning sites vary by

clone and include the following: 5'-EcoRV-XmiI/XhoI-3',

5'-EcoRV-XmiI/NotI-3', EcoRV (TA cloned, non-directional).

For information about which gene each clone represents,

please visit our anonymous ftp site at

ORIGIN

ftp://image.llnl.gov/image/reatrayed_plates/IRB1.presv.dat
a Note: this is a NIH_MGC Library."

Query Match 66.6%; Score 690.8; DB 7; Length 824;
Best Local Similarity 97.1%; Pred. No. 3.6e-165;
Matches 736; Conservative 0; Mismatches 17; Indels 5; Gaps 3;

QY 1 TCAGGAATGATGCGCTTTGGCACAATATTAATATATTTCTGTGTGAAAAACAATG 60
Db 14 TCAGGAATGATGCGCTTTGGCACAATATTAATATATTTCTGTGTGAAAAACAATG 73
QY 61 TCAATATGATGCGGATGCTTCCCTGACATTTAATAGTGTCTCATATTTGACACATC 120
Db 74 TCAATATGATGCGGATGCTTCCCTGACATTTAATAGTGTCTCATATTTGACACATC 133
QY 121 GTTGCAATTCGATATGTTATTTCTATATACACTTCAACAACCTTATACCCCAACA 180
Db 134 GTTGCAATTCGATATGTTATTTCTATATACACTTCAACAACCTTATACCCCAACA 193
QY 181 AATTGCTCATTCATTCATGAGGACCATGAGACCTTTCTGTGGGATGCTGTGATGCT 240
Db 194 AATTGCTCATTCATTCATGAGGACCATGAGACCTTTCTGTGGGATGCTGTGATGCT 253
QY 241 TACAGTATGATGATGATGCTGACGACATGTTGATATTTGAGAAAGCTTCTGTAAT 300
Db 254 TACAGTATGATGATGATGCTGACGACATGTTGATATTTGAGAAAGCTTCTGTAAT 313
QY 301 CACACAGACCGACATTTATGCTGACGCTGACCTTCCATTTTCTTCTTCAATCTCC 360
Db 314 CACACAGACCGACATTTATGCTGACGCTGACCTTCCATTTTCTTCTTCAATCTCC 373
QY 361 ATTGACCGCTATGCTGTGTGTGATGACATGATATTAAGCAAGATGATATCTTG 420
Db 374 ATTGACCGCTATGCTGTGTGTGATGACATGATATTAAGCAAGATGATATCTTG 433
QY 421 GTTATTTGTGATGATCTTCAATTAATAGTGAAGTGTCCCTGCTGTTTTTGCAATTTG 480
Db 434 GTTATTTGTGATGATCTTCAATTAATAGTGAAGTGTCCCTGCTGTTTTTGCAATTTG 493
QY 481 ATCTTTCTGAGCTAACTTCAAGGCGCTGAAGAGATATATATACAAACATGTCACCTG 540
Db 494 ATCTTTCTGAGCTAACTTCAAGGCGCTGAAGAGATATATATACAAACATGTCACCTG 553
QY 541 AGAGAGAGTGC--TCTGCTCTTTAGCAAAATATCTGGGTACTGACCTTTATGACTTC 599
Db 554 AGAGAGAGTGC--TCTGCTCTTTAGCAAAATATCTGGGTACTGACCTTTATGACTTC 613
QY 600 TTTTATATACCTGATCTATTTATGTTATGTCCTATTAAGATATATTTATTCGCTAA 659
Db 614 TTTTATATACCTGATCTATTTATGTTATGTCCTATTAAGATATATTTATTCGCTAA 673
QY 660 AGACAGGCAAGATTAATAGTAGTGC--AATCAGAACTCCAAATTTG--ATTGAAATGA 718
Db 674 AGACAGGCAAGATTAATAGTAGTGC--AATCAGAACTCCAAATTTG--ATTGAAATGA 733
QY 719 AAAATGAAATTT--CACAAAGCAAGAAAGAAAGCT 753
Db 734 AAAATGAAATTTTCCCAAAACCAAGAAAGAAAGCT 771

RESULT 4

AY406936

LOCUS AY406936

DEFINITION Pan troglodytes HCM723 gene, VIRUAL TRANSCRIPT, partial sequence,

genomic survey sequence.

ACCESSION AY406936

VERSION AY406936.1 GI:39762907

KEYWORDS GSS.

SOURCE Pan troglodytes (chimpanzee)

ORGANISM Pan troglodytes

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.

REFERENCE 1 (bases 1 to 675)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 675)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES
source
Location/Qualifiers
1..675
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..5675
/locus_tag="HMC2723"
ORIGIN
Query Match 63.3%; Score 657; DB 9; Length 675;
Best Local Similarity 97.8%; Pred. No.1.5e-156;
Matches 660; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 352 TTGATCTCCATTGACCGCTACTAGTGTGTGTGATCCAGATATATAAGCAAGATG 411
DB 1 TTGATCTCCNNNNAACGCTACTANNNNNTGTGTGATCCAGATATATAAGCAAGATC 60
QY 412 AATATCTGGTATTGTTGTGTGTGTGATCTTCAATTAAGTGAAGTGTCCCTGCTTTTGCA 471
DB 61 AATATCTGGTATTGTTGTGTGTGTGATCTTCAATTAAGTGAAGTGTCCCTGCTTTTGCA 120
QY 472 TTGGAATGATCTTCTGTGAGCTTAACTCAAGGCGTGAAGAGATATATTACAAAT 531
DB 121 TTGGAATGATCTTCTGTGAGCTTAACTCAAGGCGTGAAGAGATATATTACAAAT 180
QY 532 GTTCACTGAGAGAGGTTGCTGTCTCTTTTCAAAATATCTGGGCTACTGACCTTT 591
DB 181 GTTCACTGAGAGAGGTTGCTGTCTCTTTTCAAAATATCTGGGCTACTGACCTTT 240
QY 592 ATGACTCTCTTTTATATACCTGATCTATTATGTTATGTCATATTAAGATATATCTT 651
DB 241 ATGACTCTCTTTTATATACCTGATCTATTATGTTATGTCATATTAAGATATATCTT 300
QY 652 ATGCTTAAAGACAGGCAAGATTAATTAAGTGCATCAGAAAGCTCAATTTGGAATG 711
DB 301 ATGCTTAAAGACAGGCAAGATTAATTAAGTGCATCAGAAAGCTCAATTTGGAATG 360
QY 712 GAATGAAGAAATGGAATTTCAAAAGCAAGAAAGAAAGCTGTGAAGACATTGGGAAAT 771
DB 361 GAATGAAGAAATGGAATTTCAAAAGCAAGAAAGAAAGCTGTGAAGACATTGGGAAAT 420
QY 772 GTGATGGAGATTTCTAATATGCTGTGCTCTTTCTTATCTGTACAGTATGGAACCT 831
DB 421 GTGATGGAGATTTCTAATATGCTGTGCTCTTTCTTATCTGTACAGTATGGAACCT 480
QY 832 TTTCTTCACTACATTATTCACCTACTTGAATGATGTTGATTTGGTTGGCTACTTG 891
DB 481 TTTCTTCACTACATTATTCACCTACTTGAATGATGTTGATTTGGTTGGCTACTTG 540
QY 892 AACTCTACATTTAATCCAAATGTTTATGCAATTTTCAATCTGTTTGAAGAAAGCACTG 951
DB 541 AACTCTACATTTAATCCAAATGTTTATGCAATTTTCAATCTGTTTGAAGAAAGCACTG 600
QY 952 AAGATGATGCTGTTGGTAAATTTTCCAAAAGATTCATCAGGTGTAATATTGTTT 1011

DB 601 AAGATGATGCTGTTGGTAAATTTTCCAAAAGATTCATCAGGTGTAATATTGTTT 660
QY 1012 GAATGAGTCACTAG 1026
DB 661 GAATGAGTCACTAG 675
RESULT 5
AY406937
LOCUS
DEFINITION Mus musculus HMC2723 gene, VIRtual TRANSCRIPT, partial sequence,
AY406937
ACCESSION
VERSION AY406937.1 GI:39762908
KEYWORDS GSS.
SOURCE
ORGANISM Mus musculus (house mouse)
REFERENCE 1 (bases 1 to 927)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 927)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES
source
Location/Qualifiers
1..927
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>927
/locus_tag="HMC2723"
ORIGIN
Query Match 58.0%; Score 601.8; DB 9; Length 927;
Best Local Similarity 78.7%; Pred. No.1.9e-142;
Matches 733; Conservative 0; Mismatches 192; Indels 6; Gaps 1;
QY 94 ATGCTCTCATATATTTGACCACTGCTGGCAATCTGATAGTATGTTTCTATATCA 153
DB 1 ATGCTCTCATATATCTGGCACTGCTGGCACTGATAGTATGTTTCTATATCA 60
QY 154 CACTTCAACAACCTCATACCCCAACAAATGGCTCATTTCAATTCATGCGCCCTGAC 213
DB 61 CATTTCACCAACTTCATACACCCCAACAAATGGCTCTTCACTCAATGCGCCCTGAC 120
QY 214 TTTCTTGGGGGTGTCTGTGTCATAGCTTAACATATGATGAGATCTGTGAGCACTGTTG 273
DB 121 TTTCTTGGGGGTGTCTGTGTCATAGCTTAACATATGATGAGATCTGTGAGCACTGTTG 180
QY 274 TATTTTGGAGAGTCTTCTGTAAATTTCAACAGACCGACATTAAGTGAAGTCAAGC 333
DB 181 TATTTTGGAGAGTCTTCTGTAAATTTCAACAGACCGACATTAAGTGAAGTCAAGC 240
QY 334 TCCATTTTCCATTTGCTTTTCACTCCATGACCGCTACTAGTGTGATGATGCACTG 393
DB 241 TCCATTTTCCATTTGCTTTTCACTCCATGACCGCTACTAGTGTGATGATGCACTG 300
QY 394 AGATATTAAGCCAGATGATATCTTGGTATTTGTGTGATGATCTTCAATTAAGTGAAGT 453

```

Db 301 AGATACAAAGCCAAAGTCAATATCTCCACTATTTCTTGATGATCTCGTTAGTGGAGC 360
QY 454 GTCCCTGCTGTTTTCATTTGGAAATGATCTTTCTGGAGCTAACTCAAGGCGCTGAA 513
Db 361 CTTCCTGCTGTTTTCATTTGGAAATGATCTTTCTGGAGCTAACTCAAGGCGCTGAA 420
QY 514 GAGATATATATCAAAATGATGATCTTCAGACAGGAGGTGCTGCTCTTTAGCAAAATA 573
Db 421 GAGCTGTATCGACATGACAGTCAAGCAGCAGCGGCGGCTGTTCTCTTTTGTAAAGTA 480
QY 574 TCTGGGGATCTGACCTTTATGACCTTTTATATATACCTGATCTATATATGTTATGTC 633
Db 481 TCTGGGGATCTGACCTTTATGACCTTTTATATATACCTGATCTATATATGTTATGTC 540
QY 634 TATTACAGATATATCTTATGCTTAAAGACAGGCAAGTATATATATGATGCCATCAG 693
Db 541 TACTATAGATATATTTCTATGCTTAAAGACAGGCAAGTATATATGATGCCATCAG 594
QY 694 AAGCTCAAAATTTGATTTGGAATGAAATGGAATTTCAAGCAAGCAAGCAAGCAAGT 753
Db 555 AATGTTCAAGTTGATTTGGAATGGAATGGAATTTCAAGCAAGCAAGCAAGCAAGT 654
QY 754 GTGAAGACATTTGGGATTTGATGATGAGATTTTCTAATATGCTGTCCTTTCTTTATC 813
Db 655 GCGAAGACCTTATGAGGATCATGATGAGGCGTTTCTCTGATATGCTGTCCTTTCTTC 714
QY 814 TGTACAGTCATGACCCCTTTTCTTCACTACATATTTCCACTGATGATGATGTC 873
Db 715 TGACAGGTCCTGACCCCTTTTCTGAGCTATGTTATCCACCTCTCTGATATGACACAG 774
QY 874 ATTGCTTGGCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 933
Db 775 TATTGCTTGGCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 834
QY 934 TGGTTAGAAAAGCACTGAAGATGATGCTGTTTGTAAATTTTCCAAAGATTCATCC 993
Db 835 TGGTTAGAAAAGCACTGAAGATGATGCTGTTTGTAAATTTTCCAAAGATTCATCC 994
QY 994 AGGTGAATTTTATTTTGAATTTGATGAT 1024
Db 895 AGGTGAATTTTATTTTGTAAAGCAATTCAT 925

```

RESULT 6
 LOCUS CO957870/c 789 bp mRNA linear EST 17-AUG-2004
 DEFINITION AGENCOURT 30842301 NIH MGC_146 Homo sapiens cDNA clone
 IMAGE:7389847 3', mRNA sequence.
 ACCESSION CO957870
 VERSION CO957870.1 GI:51322452
 KEYWORDS EST.
 ORGANISM Homo sapiens (human)
 SOURCE
 Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NIH-MGC <http://mgi.mgi.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Guthrie cDNA Resource Center
 cDNA Library Preparation: Guthrie cDNA Resource Center
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: IRB18 row: 9 column: 05

FEATURES
 source
 High quality sequence start: 24
 High quality sequence stop: 579.
 Location/Qualifiers
 1..789

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:7389847"
 /issue_type="mixed"
 /lab_host="DH10B (TI-phage-resistant)"
 /clone_id="NIH_MGC_146"
 /note="Vector: pCDNA3.1; Site 1: multiple; Site 2:
 multiple; ORF's were PCR-amplified (from IMAGE clones or
 from commercially available cDNA libraries) and cloned by
 the Guthrie cDNA Resource Center (www.guthrie.org/cDNA)
 into pCDNA3.1. For specific information on cloning sites
 (which vary by clone), please refer to the Guthrie
 website, using the Guthrie ID given in the file
ftp://image.llnl.gov/image_rearayed_plates/IRB18.presv.dat
 a. Note: this is a NIH_MGC Library."

Query Match 53.3%; Score 553.2; DB 7; Length 789;
 Best Local Similarity 99.1%; Pred. No. 4.7e-130;
 Matches 577; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

```

QY 422 GATGATCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 491
Db 614 GATGATCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 557
QY 492 GCTAACTTCAAAAGCGCTGAAGAGATATATTACAAATGTTCACTGACAGAGAGTTG 551
Db 556 GCTAACTTCAAAAGCGCTGAAGAGATATATTACAAATGTTCACTGACAGAGAGTTG 497
QY 552 CTGCTCTTCTTTAGCAAAATATCTGGGTACTGACCTTTATGACTCTTTTATATACC 611
Db 496 CTGCTCTTCTTTAGCAAAATATCTGGGTACTGACCTTTATGACTCTTTTATATACC 437
QY 612 TGGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 671
Db 436 TGGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 377
QY 672 ATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 731
Db 376 ATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 317
QY 732 ACAAGCAAGAAAGAAAGCTGTGAAGACATTTGGGATTTGATGAGAGTTTCTTAT 791
Db 316 ACAAGCAAGAAAGAAAGCTGTGAAGACATTTGGGATTTGATGAGAGTTTCTTAT 257
QY 792 ATGCTGTGCTCTTTCTTTATCTGTACAGTATGAGACCTTTTCTTCACTATATTC 851
Db 256 ATGCTGTGCTCTTTCTTTATCTGTACAGTATGAGACCTTTTCTTCACTATATTC 197
QY 852 ACTTACTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 911
Db 196 ACTTACTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 137
QY 912 GATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 971
Db 136 GATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 77
QY 972 AATTTTCCAAAGATTTCAATCCAGGTGTAATTTATTTTGA 1013
Db 76 AATTTTCCAAAGATTTCAATCCAGGTGTAATTTATTTTGA 35

```

RESULT 7
 LOCUS CO957854 684 bp mRNA linear EST 17-AUG-2004
 DEFINITION AGENCOURT 30842669 NIH_MGC_146 Homo sapiens cDNA clone
 IMAGE:7389847 5', mRNA sequence.
 ACCESSION CO957854

VERSION C0957854.1 GI:51322436
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-remail.nih.gov
Tissue Procurement: Guthrie cDNA Resource Center
cDNA Library Preparation: Guthrie cDNA Resource Center
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: IRB18 row: 9 column: 05
High quality sequence stop: 460.
Location/Qualifiers

FEATURES
source
1.684
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7389847"
/issue_type="mixed"
/lab_host="DH10B (T1-phage-resistant)"
/note="Vector: pCDNA3.1; Site 1: multiple; Site 2: multiple; ORF's were PCR-amplified (from IMAGE clones or from commercially available cDNA libraries) and cloned by the Guthrie cDNA Resource Center (www.guthrie.org/cDNA) into pCDNA3.1. For specific information on cloning sites (which vary by clone), please refer to the Guthrie website, using the Guthrie ID given in the file ftp://image.llnl.gov/image.rearrayed_plates/IRBF.presv.dat
a Note: this is a NIH_MGC library."

ORIGIN

Query Match 49.0%; Score 508.4; DB 7; Length 684;
Best Local Similarity 96.9%; Pred. No. 1.2e-118;
Matches 526; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

```
7 ATGATGCCCTTGGCACATATATTAATTTCTGTGGAAGAAACAGCGGCAAT 66
2 ATGATGCCCTTGGCACATATATTAATTTCTGTGGAAGAAACAGCGGCAAT 61
67 GATGCCCTTGGCACATATTAATTTCTGTGGAAGAAACAGCGGCAAT 126
62 GATGCCCTTGGCACATATTAATTTCTGTGGAAGAAACAGCGGCAAT 121
127 AATCGATAGTATTTGTTCTATATCACACTTCAACCACTTACCCCAAAATTGG 186
122 AATCGATAGTATTTGTTCTATATCACACTTCAACCACTTACCCCAAAATTGG 181
187 CTCATTCATTCACAGCCAGCTGTGAGCTTTCTTCTGGGGTGTCTGATCAGT 246
182 CTCATTCATTCACAGCCAGCTGTGAGCTTTCTTCTGGGGTGTCTGATCAGT 241
247 ATGGGAGATCTGCTGAGCACTGTGGATTTTGGAGAAGCTTCTGTAATTCACACA 306
242 ATGGGAGATCTGCTGAGCACTGTGGATTTTGGAGAAGCTTCTGTAATTCACACA 301
307 AGCACCAGCATTTAGCTGAGCTCAGCTCCATTTTCATTGTTCACTCCATTGAC 366
302 AGCACCAGCATTTAGCTGAGCTCAGCTCCATTTTCATTGTTCACTCCATTGAC 361
367 CGCTACTATGCTGTGTGATCCAGTGAATATAAGCAAGATGATATCTGTTATT 426
```

```
|||||
362 CCTTACTATGCTGTGTGATCCAGTGAATATAAGCAAGATGATATCTGTTATT 421
427 TGTGATGATCTTTCATTAGTTGAGTGTCCCTGTGTTTGGATTGATGATCTTT 486
422 TGTGATGATCTTTCATTAGTTGAGTGTCCCTGTGTTTGGATTGATGATCTTT 481
487 CTGGAGCTTAACTTCAAGGCGCTGAGAGATATATTCAAAACATGTTCACTGCAGAGA 546
482 CTGGAGCTTAACTTCAAGGCGCTGAGAGATATATTCAAAAC-TGTTTCTTTAGAGA 540
QY 547 GGT 549
Db 541 GGT 543
```

RESULT 8
CN838940/c
LOCUS CN838940
DEFINITION AGNCOURT.15864245 NIH_MGC.145 Homo sapiens cDNA clone
IMAGE:30706625 3', mRNA sequence.
ACCESSION CN838940
VERSION CN838940.1 GI:47944595
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-remail.nih.gov
Tissue Procurement: GPCR Consortium
cDNA Library Preparation: GPCR Consortium
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: IRB14 row: f column: 06
High quality sequence stop: 689.
Location/Qualifiers

FEATURES

source

1.929
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30706625"
/issue_type="mixed"
/lab_host="DH10B"
/note="Vector: pCDNA3.1; Site 1: varies by clone; Site 2: varies by clone; ORF's were PCR-amplified and cloned into pCDNA3.1 by the GPCR Consortium. Cloning sites vary by clone and include the following: 5'-EcoRV-XmiI/XhoI-3', 5'-EcoRV-XmiI/NotI-3', EcoRV (TA cloned, non-directional). For information about which gene each clones represents, please visit our anonymous ftp site at ftp://image.llnl.gov/image/rearrayed_plates/IRB1.presv.dat
a Note: this is a NIH_MGC library."

ORIGIN

Query Match 41.6%; Score 431.4; DB 7; Length 929;
Best Local Similarity 98.5%; Pred. No. 5.9e-99;
Matches 446; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

```
587 CCTTATGACTTC-TTTTATATACCTGATCTATTATGATGCTTATTAACAGATA 645
Db 752 CCTTATGACTTCTTTTATATACCTGAATCTATTATGATGCTTATTAACAGATA 693
```

QY 646 TATCTATCGCTAAAGACAGGCAAGATTAAATGATGATCCCAATCAGAACTCCAAATT 705
| | | | |
Db 692 TATCTATCGCTAAAGACAGGCAAGATTAAATGATGATCCCAATCAGAACTCCAAATT 633
| | | | |
QY 706 GGATTGGAATGAAATGAAATTTTCACAAAGCAAGAAAGAAAGCTGTGAACATTTG 765
| | | | |
Db 632 GGATTGGAATGAAATGAAATTTTCACAAAGCAAGAAAGAAAGCTGTGAACATTTG 573
| | | | |
QY 766 GGGATTGATGGAGTTTCTTAATATGCTGTGCTCTTTATCTGTACATCATG 825
| | | | |
Db 572 GGGATTGATGGAGTTTCTTAATATGCTGTGCTCTTTATCTGTACATCATG 513
| | | | |
QY 826 GACCTTTTCTTCACTACATTAATCCACCTACTTGAATGATGTGTGTTGGTGGC 885
| | | | |
Db 512 GACCTTTTCTTCACTACATTAATCCACCTACTTGAATGATGTGTGTTGGTGGC 453
| | | | |
QY 886 TACTTGAACCTACATTTAATCAATGTTTATGCAATTTTCTATCTGTGTTAGAAA 945
| | | | |
Db 452 TACTTGAACCTACATTTAATCAATGTTTATGCAATTTTCTATCTGTGTTAGAAA 393
| | | | |
QY 946 GCACGAGAGATGATGCTGTTTGTGTAATAATTTCCAAAAGATTCATCCAGGTGAATTA 1005
| | | | |
Db 392 GCACGAGAGATGATGCTGTTTGTGTAATAATTTCCAAAAGATTCATCCAGGTGAATTA 333
| | | | |
QY 1006 TTTTGAATGATGATCAATGATTAATTTATTT 1038
| | | | |
Db 332 TTTTGAATGATGATCAATGATTAATTTATTT 300
| | | | |

RESULT 9
CN832787/c 941 bp mRNA linear EST 02-JUN-2004
LOCUS CN832787
DEFINITION AGECOIRT.15864261 NIH MGC_145 Homo sapiens cDNA clone
IMAGE:30706626 3', mRNA sequence.
CN832787
ACCESSION CN832787
VERSION
KEYWORDS EST.
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 941)
NIH-MGC http://mgi.mc.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgraphs-remail.nih.gov
Tissue Procurement: GPCR Consortium
cDNA Library Preparation: GPCR Consortium
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
Place: IRB14 row: E column: 07
High quality sequence scop: 646.
Location/Qualifiers
1..941
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30706626"
/isue_type="mixed"
/lab_host="DH10B"
/clone_lib="NIH MGC 145"
/note="Vector: pCDNA3.1, Site 1: varies by clone; Site 2:
varies by clone; ORFs were PCR-amplified and cloned into
pCDNA3.1 by the GPCR Consortium. Cloning sites vary by
clone and include the following: 5'-EcoRV-XbaI/XhoI-3',

5'-EcoRV-XbaI/XhoI-3', EcoRV (TA cloned, non-directional).
For information about which gene each clones represents,
please visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearrayed_plates/IRB1.presv.dat
a Note: this is a NIH_MGC library."

Query Match 40.2%; Score 417; DB 7; Length 941;
Best Local Similarity 94.3%; Pred. No. 2,8e-95;
Matches 465; Conservative 0; Mismatches 25; Indels 3; Gaps 3;

QY 549 TTGCTGTCTCTTTAGCAAAATATCTGGGAGTCACTTATGAC-TTCTTTTATA 607
| | | | |
Db 795 TTTTCTTCTTATAGCAAAATATATGCGGAGTCACTTATGACCTTTCTTTTATA 736
| | | | |
QY 608 TACTCGAT-CTATTATGTATGTCTTATACAAATATCTTATGCTTAAACAG 666
| | | | |
Db 735 TCCCTGATGCTCTTATGTATGTCTTATACAAATATCTTATGCTTAAACAG 676
| | | | |
QY 667 GCAAGATTAAATGATGATGCCAATCAGAAAGCTCCAAATTGATGGAATGAAATGGA 726
| | | | |
Db 675 GCAAGATTAAATGATGATGCCAATCAGAAAGCTCCAAATTGATGGAATGAAATGGA 616
| | | | |
QY 727 ATTTCACAAAGCAAGAAAGAAAGC-TGTGAAGACATTGGGATGTGATGGAATTT 785
| | | | |
Db 615 ATTTCACAAAGCAAGAAAGAAAGC-TGTGAAGACATTGGGATGTGATGGAATTT 556
| | | | |
QY 786 CCTAATATGCTGTGCTCTTTCTTATCTGTACATGACCTTTTCTTCACTACAT 845
| | | | |
Db 555 CCTAATATGCTGTGCTCTTTCTTATCTGTACATGACCTTTTCTTCACTACAT 496
| | | | |
QY 846 TATTCACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 905
| | | | |
Db 435 TATTCACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 436
| | | | |
QY 906 TCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 965
| | | | |
Db 435 TCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 376
| | | | |
QY 966 TGTGTAATTTTCCAAAAGATTCATCCAGGTGAATTTTGGATGATGATGATGAT 1025
| | | | |
Db 375 TGTGTAATTTTCCAAAAGATTCATCCAGGTGAATTTTGGATGATGATGATGAT 316
| | | | |
QY 1026 GAATTTATTAATT 1038
| | | | |
Db 315 GAATTTATTAATT 303
| | | | |

RESULT 10
CNS04073/c 1016 bp DNA linear GSS 01-SEP-2000
LOCUS CNS04073/c
DEFINITION Tetradon nigroviridis genome survey sequence PUC-Orl end of clone
129C03 of library G from Tetradon nigroviridis, genomic survey
sequence.
ACCESSION AL302248.1 GI:8180344
VERSION
KEYWORDS GSS: genome survey sequence.
SOURCE
ORGANISM Tetradon nigroviridis
Tetradon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percormorpha; Tetradontiformes;
Tetradontidae; Tetradontidae; Tetradon.
1
Roest Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fzames,C., Wincker,P., Brothier,P., Queller,F.,
Saurin,W. and Weissenbach,J.
Estimate of human gene number provided by genome-wide analysis
using Tetradon nigroviridis DNA sequence
Nat.Genet. 25 (2), 235-238 (2000)
JOURNAL
MEDLINE
PUBMED
20296633
10635645
REFERENCE

AUTHORS
Roest Crolius,H., Jallion,O., Dasilva,C., Ozouf-Costaz,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
TITLE
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL
Genome Res. 10 (7), 939-949 (2000)
MEDLINE
20359837
10899143
3 (bases 1 to 1016)
REFERENCE
Genoscope.
AUTHORS
Direct Submission
TITLE
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
JOURNAL
- Web : www.genoscope.cns.fr
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
Location/Qualifiers
1..1016
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone_1="129C03"
/clone_11b="Tg"
/note="Genoscope sequence ID : C0BGI29AB02SP1-end :
PUC-Or1"
ORIGIN
Query Match 25.3%; Score 262.8; DB 9; Length 1016;
Best Local Similarity 55.2%; Pred. No. 6.4e-56;
Matches 544; Conservative 0; Mismatches 423; Indels 18; Gaps 1;
QY 34 AATATTTCTGTGTGAAAAACAACACTGTGCAATGATGTCGGTCTCCCTGACAGTTTA 93
DB 967 AACATTTCTGTGTGAAAGACCGCTATCACTGCCCTTTCGTGATCCTTTACATGGTT 908
QY 94 ATGGGCTCATTAATTTGACCAACTGTTGGCAATCTGATAGTATTTGTTATATCA 153
DB 907 TTAGGGGTATGTCATGTCGTCGTGTGGAACCTTTGGTCACGTTTCTATCATC 848
QY 154 CACTCAACAACACTCATACCCCAAAATGGCTCATTCATTCATGAGCCACTGTGAC 213
DB 847 TATTTTAAGCACTCACTCAACCACTACCTGTTGCTCTCTGCTGTGTCTGAC 788
QY 214 TTTCTTCTGGGGTGTGTGTCATGCTTACAGTATGTGATGTGTGAGCACTGTGG 273
DB 787 TTTCTCTGGGGTGTGTGTCATGTTACCAATATGATTCAGACATGAAACCTGCTGG 728
QY 274 TATTTTGAAGAAGTTTCTGTAAATTCACACAGCAGCAGCATTTAGCTGACGCC 333
DB 727 TATTTTGAAGACATTTTGGCAAGTTCATTGAGTTCTGATGTGATGTTGTGACAGCA 668
QY 334 TCCATTTTCAATTTGCTTTCATCTCCATGACCGCTACTACTGCTGTGTGATGCCACTG 393
DB 667 TCCATTAATTAATTCGCTTTTAATATCATTTGATGATACATTCGGGATTCATCTCTTA 608
QY 394 AGATTAAGCAAGCAATATCTTGGTATTTGTGTGATGATCTTCATTAGTTGAGT 453
DB 607 GTGATTAAGAAAAAATGACTGTGATGCTGTTTGTGTGATGATTTTGTGACGTGAGT 548
QY 454 GTCCCTGTCTTTTGTGATTTGGAATGATCTTCTGAGAGCTTAACCTTCAAAAGCGCTGAA 513
DB 547 GTTTCAGGTGTGTGAGTTTGGATGATTTTAAAGCTGAACATTTGGGGAATTGAA 488
QY 514 GAGATTAATTAACAACATGTTCACTGCAAGAGAGTGTCTGTCTCTTTCAAAATA 573
DB 487 GATTTTCAACAACATTTGTTGATGGAAGAGATGATTTTGTGACAGAGGTCCG 428
QY 574 TCTGGGGTACGACCTTATGACTCTTTTATATACCTGATCTATTAATGTTATGTGTC 633
DB 427 TCAAGTACAGTCTCATGATCTCTCTTTTATATTCAGGAATTAATCATGATCATGTTTA 368

QY 634 TATTACAGAAATATCTTATCCGTTAAAGACAGCAAGATTATATGATGCCAATCAG 693
DB 367 TAGCGAAGATTTCTTTGTGTGCAAAAGACAAATTCCTCAGATTCAGAAATCAAGCAGC 308
QY 694 AGCTCCAAATTTGATTTGGAATGAAAATGGAATTTCACAAAGCAAAAGGAAGCT 753
DB 307 ATG-----ACATCAGCCAGAGATTCAGATTAAGAAAAACAGCAAAAGCC 266
QY 754 GTGAAGACATTTGGGATTTGTGATGGAGTTTCTTAATATGCTGTGCTTCTTATC 813
DB 265 ACTAAACCTTGTGTGATCATGAGACATTTCTTTATGCTGAGCTTCAGTTTCTC 206
QY 814 TGTAAAGTCATGACCTTTTCTTCACTACATTAATTCACCTACTTTGATGATGTG 873
DB 205 TGTTCATCATTTGATTCCTTTTTCAGCTACCAACCAACAGATGATTTGAAACATTC 146
QY 874 ATTTGTTGGCTACTTGAACCTTACATTAATTAATGATTTTATGATTTTCTATCT 933
DB 145 CTATGCTAGTTATATGATTTTCACTAAACCTCTCATTTATGCAATTTTTCACACC 86
QY 934 TGGTTTGAAGAAAGCACTGAAGATGATGCTGTTGTGTAATTTTCCAAAAAGATTCAACC 993
DB 85 TGGTTTGAAGAGCGTTTTCCTGTTTGTCTTCAGAGAAAGCTTTAAAGAGATATGTCA 26
QY 994 AGCTGTAAATTTATTTTGGAAATGA 1018
DB 25 AACATTTCTCTTTTACTTCATTA 1
RESULT 11
CNS05JP2 1015 bp DNA linear GSS 01-SEP-2000
Tetraodon nigroviridis genome survey sequence T7 end of clone
008G08 of library A from Tetraodon nigroviridis, genomic survey
sequence.
AL340463.1 GI:8234221
VERSION
GSS: genome survey sequence.
KEYWORDS
Tetraodon nigroviridis
Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE
1
Roest Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fischer,C.,
Bernot,A., Fizames,C., Wincker,P., Brotilier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
TITLE
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
JOURNAL
Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE
20296633
10835645
2
REFERENCE
2
Roest Crolius,H., Jallion,O., Dasilva,C., Ozouf-Costaz,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
TITLE
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish tetraodon nigroviridis
JOURNAL
Genome Res. 10 (7), 939-949 (2000)
MEDLINE
20359837
10899143
3 (bases 1 to 1015)
REFERENCE
Genoscope.
AUTHORS
Direct Submission
TITLE
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
JOURNAL
- Web : www.genoscope.cns.fr
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
Location/Qualifiers

```

source
1. 1015
/organism="Tetrarodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="008608"
/clone_lib="A"
/notes="Genoscope sequence ID : C0AA008B04C1-end : T7"

Query Match      24.8%; Score 257.6; DB 9; Length 1015;
Best Local Similarity 55.9%; Pred. No. 1,4e-54;
Matches 510; Conservative 0; Mismatches 385; Indels 18; Gaps 1;

OY 34 AATATTTCTGTGTGAAAAAACAACAGTGTGAAATGATGTCGGTCTTCCCTGACAGTTTA 93
DB 98 AACATTCTGTGTGAAAAACGCTCTATTCACCTGCCATTGCGTATCATCTTTACATGTT 157
OY 94 ATGTGCTCATATTTCTGACCACTCGTTGGCAATCTGATAGTTATGTTTCTATATCA 153
DB 158 CTAGGGGNCATGTTGTGTGACGTGTGTGAAACCTTTGTGTCAGTGTTCATATTC 217
OY 154 CACTTCAACCACTTATACCCCAACAAATTTGGCTATTCATCCATGGCCACTGTGAC 213
DB 218 TATTTTAAAGACCTCCACACTCCAACTACCTGTGTGCTCTTGTGTGTGTAC 277
OY 214 TTTCTTGGGGTGTCTGTGTCAGCTTACAGTATGTGATGTCGTGACACTGTTGG 273
DB 278 TTTCTCTGGGGTGTCTGTGTCATGTATCCCATATGATTCATTTGGCAGACCTGTGG 337
OY 274 TATTTTGAAGAAAGTCTTGTGAAATTCACACAGACCCGACATTAATGCTGAGCTCAGCC 333
DB 338 TATTTTGAAGACATTTTGGCAAGATGACATNNNGATTTGCTGNAATGTTGTGACAGCA 397
OY 334 TCCATTTTCCATTTGTCTTTCATCTCCATGACCGCTACTATGCTGTGTGTATCCACTG 393
DB 398 TCCATTAATTAATCTGCTTAAATATCAATGATGATGATGAAAGCGGTATCATCTCTTA 457
OY 394 AGATATTAAGCCAGATGAAATATCTTGGTATTTGTGTGATGATCTTCAATGTTGAGT 453
DB 458 GTGTATTAAGAAAAATGTCTGTGATGTTGTTGTGTATATTTTGTACGTGGGT 517
OY 454 GTCCCTGTCTTTTGTGATTTGAAATGATCTTCTGAGACTTAACCTTCAAGCGCTGAA 513
DB 518 GTTTCAGGTGTGTGATTTGGAATGATTTTTCAGGCTGAATATTTTGGGAATGAA 577
OY 514 GAGATATATTAACAACATGTTCACTGACAGAGAGGTGCTGTCTCTTTAGCAAAATA 573
DB 578 GATTTCTCAATGACCTTTGTGTCATGTGAAAGAAATGATATTTGTTCAAGTGTATG 637
OY 574 TCTGGGTAAGTACCTTATGACCTTCTTTTATATACCTGATCTATTAATGTTATGTC 633
DB 638 TCAAGACAGTCTGCTGACGTCTCTCTTTTATATCCAGAAATATCATGATTTGTGTA 697
OY 634 TATTACAGAAATTAATCTTATCGTAAAGAACGAGAAATTAATTAATGATGATCCAAATG 693
DB 698 TACCTGAAGATTTTCTTTGTGCAAGAGACAAATTCCTCAGCATTCAGAAATCAACCTGC 757
OY 694 AAGCTCCAAATGATGGAATGAATAATGAAATTTTCAAGCAAGAAAGAAAGT 753
DB 758 ATG-----ACATCAGCAGCAATTTCAAAATTAAGCAAGCAAAAGC 799
OY 754 GTGAGACATTTGGGATTTGTGATGGAGTTTCTTAATATGCTGTGACCTTTCTTTATC 813
DB 800 ACTAATAACCTGTTGTGTATCATGAGAGGTTTCTTTATGCTGGCTCATTTTTCTC 859
OY 814 TGTACAGTCATGAGACCTTTTCTTCACTACATTAATTCACCTTAATGATGATGTTG 873
DB 860 TGTATATGTCATGATCCTTTCACTGACCTCTCAACACGACGATGATGAAACATTT 919
OY 874 ATTGTGTTGGCTACTGAACTCTACATTAATCCAAATGTTTATGCAATTTTCTATCT 933
DB 930 GGATGGGNGGTTTATTAATCTTACACAAACCCATCATTAATGCAATTTTAAACCC 979

```

```

OY 934 TGGTTTAAAGAAAG 946
DB 980 TGGTTCAAAAG 992

RESULT 12
AM813121/C
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 536)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.G.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?el=et2-CM4-ST0189-051
099-021-h04&e3=1999-10-05&e4=1)
Seq primer: puc 18 forward
High quality sequence start: 26
High quality sequence stop: 534.
Location/Qualifiers
1. 536
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="ST0189"
/notes="Organ: stomach; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORSTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match      23.9%; Score 248; DB 2; Length 536;
Best Local Similarity 98.0%; Pred. No. 3,4e-52;
Matches 251; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 783 TTTCTATATGCTGTGTCCTTTCTTATCTGACATCATGACCCCTTTCTTCACTA 842
DB 535 TTTCTATATGCTGTGTCCTTTCTTATCTGACATCATGACCCCTTTCTTCACTA 476
OY 843 CATATTCACCTACTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 902
DB 475 CATATTCACCTACTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 416
OY 903 TAATCCAAATGTTATGCAATTTTCTATCTTGTGTTAAGAAAGCACTGAAGATGATGCT 962

```


D _b	415	TATCCCATGGTTTAAGCATTTTCTACTCCTTGTTTGAAGAAAGACTGMACTATGCT	356
Q _y	963	GTTGGTAAATTTTTTCCAAAAGATTCAATCCAGCGTGAATAATTTTGGAAATGAGTTC	1022
D _b	355	GTTTGCTAATATTTTCCCANAAGATTCAATCCAGGTGAATAATTTTGGAAATGAGTTC	296
Q _y	1023	ATAGAATTTATATATTT	1038
D _b	295	ATAGAATTTATATATTT	280

RESULT 13	
CD559484/c	
LOCUS	
DEFINITION	797 bp mRNA linear EST 19-NOV-2003
ACCESSION	CD559484
	AGENCY:CD559484
	IMAGE:6971835 5', mRNA sequence.

FEATURES

ОБЪЕКТ

Query Match	20.5%	Score 213;	DB 6;	Length 797;
Best Local Similarity	57.6%	Pred. No. 3.1e-43;		

	Matches	437; Conservative	0; Mismatches	295; Indels	27; Gaps	2
QY	220	CTGGGGTGTCTGGTCA	TGCGCTTACAGATATGTGAGATCTGCTGACACTGTGTGATATTT	279		
Db	797	CTGGATTTTCAACATCAT	TGCGCATATATGATATGATCAATCGGTGGAACTGCTGTGATATTT	738		
QY	280	GGAGAGTCTCTGTTAAAT	TTCACAACAAGCACCGGCATTTATGCTGATGCTCAGCTCCATTT	339		
Db	737	GGGCTTACATTTTTCAG	ATTTATATATGTTTTCAGCTGATGCTTATGAGATTAACATCCACTT	678		
QY	340	TTCCATTTTGTCTTTCAT	CTCCATTCACCGCTACTATGCTGTGTGTGATCCACTGAGATAT	399		
Db	677	TTTCATCTTTGCTCAG	TGGCCATGATATGATTTTATGCTATATGTTACCCTACTTATTT	618		
QY	400	AAAGCCAAATGAAATAT	CTTGTTATTTGTGTGATATCTTCATTATGTGAGATGTCCCT	459		
Db	617	TCCACCAAAATACATAT	TCAGTCAATTAAGATATGTCACTTATATGTTGTGCGTCCCT	558		
QY	460	GCTGTTTTTTCATTTG	GAATATCTTCTGTGAGAGCTAACTTCMAAGCGCTGAAAGATATA	519		
Db	557	GGAGCATTTGCTTGG	GGGTGTCTTCTCAAG--GCTTATGCAATGATGAAATGAGGCG	501		
QY	520	TATTTCAAAACATGTT	CACTGCAAGAGAGGTGTCTGTCTCTTTAGCAAAATATCTGGG	579		
Db	500	TATGACATCTGGTGT	CTGTTTCCAGTTCTGCGCCAGTGAATGTTCAACAGCTATGGGGG	441		
QY	580	GTACTGACCTTTATG	ACTTCTTTTATATATACCTGGAATCTATATATGTTATGTCTATTTAC	639		
Db	440	ACGACCTTGTATTA	TGGCAGGTTTCTTCACTCTGGGCTTATGATGTGGGATTTATATGCG	381		
QY	640	AGAAATATATCTTAT	TGCTTAAAGAACAGGCAAGATTAATTTATGTAATGCCAATCAGAACTC	699		
Db	380	AAATTTTTCAGATAT	CCAGAAACATGCTCATGCCATCAATA-----	337		
QY	700	CAATTTGATTTGAAAT	GAATAATGAATTTTCACAAAGACAAGAAAGAAAGCTGTGAG	759		
Db	336	-----CTGCGAGAA	AAATCAAAATATATCAATGTAAGAAAGCAAAAAGCTGCCAAA	285		
QY	760	ACATTTGGGGAATGT	GATGGAGTTTCTTAAATATGCTGTGCGCTTTCTTTATCTGTACA	819		
Db	284	ACTTTAGAAATAGAT	GATGAGTTTCTTATATATGTTGGTTTCTGTATTTCTTCACAATT	225		
QY	820	GTCAATGACCTTTT	CTTCACTACATATATTCACACTTCTTGAAATGATGTGATATTTGG	879		
Db	224	TTATTTGATTCCTTT	TGAACCTTCTCATCTCCTGTATGTTTGTATATGCTTGAATGG	165		
QY	880	TTTGGCTACTGAACT	CTCAATTTAATCCAAATGATTTATGATGATTTTCTATCCCTGGTTT	939		
Db	164	TTTGTCTATTTTA	CTCACATGAAATCCGTTAATATATATGTTTCTTCTATCCCTGGTTT	105		
QY	940	AGAAATGACATGAA	ATGATGCTGTTTGTGTAATTTTTC	978		
Db	104	CGCAGACACTGAA	ATGATTTTGTCTATGTAATTTTTC	66		

RESULT	14
LOCUS	CD559544
DEFINITION	773 bp mRNA linear EST 26-NOV-2003
ACCESSION	AGNCOCOURT_14496987 NIH_MGC_135 Homo sapiens cDNA clone
VERSION	IMAGE:6971755 5' , mRNA sequence.
KEYWORDS	CD559544 CD559544.2 GI:38558970
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 773) NIH-MGC http://mgc.nci.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) On Jun 10, 2003 this sequence version replaced gi:31585612. Contact: Daniela S. Gerhardt, Ph.D.
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Narayan Bhat
cDNA Library Preparation: Bhat Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
<http://image.llnl.gov>
Plate: IRBK1 row: f column: 10
High quality sequence stop: 741.
Location/Qualifiers

FEATURES
source

```
1..773
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5971759"
/tissue_type="mixed"
/lab_host="DH5A (T1 phage-resistant)"
/clone_1lb="NIH_MGC_195"
/notes="Vector: pDNR-Dual; Site_1: loxp-Sall; Site_2:
loxP-HindIII; Clones from this library have been
PCR-amplified using gene-specific primers to contain the
complete open reading frame (based on known gene sequences
available from NCBI's RefSeq). Template for PCR is cDNA
derived from either pooled cytoplasmic polyA RNA from 30
cells lines or pooled total RNA from 10 different tissues
(from BD Biosciences/Clontech and Washington University).
PCR products are directionally cloned into the loxp sites
of the pDNR-Dual vector. Library constructed by Dr.
Narayan Bhat, Earl Bere III and Hongling Liao (Gene
Expression Laboratory, Research Technology Program, SAIC
Frederick, NCI-Frederick, Frederick, MD 21702). For
information on which gene each clone represents, please
visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearranged_plates/IRBK.presv.dat
a Note: this is a NIH_MGC library."
```

ORIGIN

Query Match 20.1%; Score 208.4; DB 6; Length 773;
Best Local Similarity 56.8%; Pred. No. 4,6e-42;
Matches 423; Conservative 0; Mismatches 317; Indels 5; Gaps 2;

```
14 CTTTTCACCAATATTAATTAATTTCTGTGTGAAACCAACTGCTCAATGATGTC 73
34 CCAAGTTGCCAAATTTGTAATAATTCCTGTCCTCCACCAACCGCTTTTCATGTC 93
74 GTGCTTCCCTGACAGTTAATGATGCTCATTAATCTGACACACTCGTTGCCAATCTGA 133
94 AGGTGATATGTAATTCGTTATGACGAGCCATGATTA--TCATTAATCGAATCTTG 151
134 TAGTAAATGTTCTATATACACTTCAACCACTTCATACCCCAACAAATGGCTCATTC 193
152 TTAATATGTTTCCATATCGCATTTCAACAGCTTCACCTCCCAACAACTTTCATGATCC 211
194 ATTCCATGGCCACTGTGACCTTTCTTCTGGGCTGTGCTCATGCTTACATATGCTGA 253
212 TCTCATAGCAACCAACGACCTTCTGCTGGGTTTGTCTATTAATGACATACACATATGC 271
254 GATGCTGACGACCTGTTGTAATTTTGAGAAAGCTTCTGTAAATTCACACAAAGACCG 313
272 GATCAGTGAAGAGTGTCTGTACTTTGGGGAATGGCTTTGTAATTCACACAAAGCTTGG 331
314 ACATTAATGCTGAGCTCAGCCCTCCATTTTCATTTGCTTCATCTTCATGACCGCTACT 373
332 ACATGATGCTCAGACTGACCTCCATTTTCCACTCTGTTCCATTTGCTATGACCGATTTT 391
374 ATGCTGTGTGATCCACTGAGATATAAGCCAAAGTGAATATCTTGGTTATTTGTGTA 433
392 ATGCGGTGTGTTACCTTTACATTAACACCAACAAATGACGAATCCACCATTAAGCAAC 451
```

```
QY 434 TGATCTCATTAAGTTGAGATGTCCTGCTGTTTTCATTTGGAATGATCTTTCTGAGC 493
DB 452 TGCTGGACATTTTGTGTCAGATTCCTGCTCTTTTCTTTTGGTTAGTTATCTGAGG 511
QY 494 TAACTTCAAAAGCGCCTGAAGATATTTACAAACATGTTCACTGACGAGAGGTGCT 553
DB 512 CCGAGT---TTCCCGTATGACGACCTTAAGATCTTGTGCTGCTTCAATTTCTGG 568
QY 554 CTGCTTCTTTAGCAAAATATCTGGGTAAGACCTTATGACTTCTTTTATATACCTG 613
DB 569 CCTTACTTTCAACAAATTTCTGGGACAAATTTGTTCACTACATGTTTCTTTACCCCTG 628
QY 614 GATCTATTATGTTATGTCATTAATACAGATATATCTTATCGCTTAAGACGAGAGAT 673
DB 629 GCTTCATCATGTTGTTGTTATGACCAAAATCTTATGCTTCCAAACGATGCTCGAG 688
QY 674 TAAATGATGATCCCATATGAAAGCTCCAAATTTGATTTGGAATGAAAATGGAATTTTCA 733
DB 689 TCATCAGCAGATGTGCTTAAACACAAAGGGGCGAGTAAACACCTATCCNAGAAAAG 748
QY 734 AAAGCAAGAAAGCAAGACGTGGA 758
DB 749 GACAGAAAGCAGCGACCTTGATA 773
```

RESULT 15
CNS04VLK
LOCUS
DEFINITION
CNS04VLK 1071 bp DNA linear GSS 01-SEP-2000
Tetradodon nigroviridis genome survey sequence T3 end of clone
035005 of library A from Tetradodon nigroviridis, genomic survey
sequence.

ACCESSION
AL030923.1 GI:8216172
VERSION
GSS: genome survey sequence.
KEYWORDS
Tetradodon nigroviridis
Tetradodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetradodon.

REFERENCE
AUTHORS
Roest Croollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Winck,P., Brotlier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
Roest Croollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetradodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)
20359837
10835645

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
Roest Croollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetradodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)
20359837
10835645

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
Roest Croollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetradodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)
20359837
10835645

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
Roest Croollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetradodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)
20359837
10835645

FEATURES
source
1..1071
/organism="Tetradodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2005, 02:33:13 / Search time 74 Seconds
(without alignments)
1771.781 Million cell updates/sec

Title: US-09-980-145-6
Perfect score: 1806
Sequence: 1 MNPFCNIIINISCVNNSN.....FGKIFKDSRCKLFLKSS 339

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp19808:*
2: geneseqp19908:*
3: geneseqp20008:*
4: geneseqp20018:*
5: geneseqp20028:*
6: geneseqp20038:*
7: geneseqp20038:*
8: geneseqp20048:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1806	100.0	339	4	AAB49232	Aab49232 Human SNO
2	1806	100.0	339	4	AAG80973	Aag80973 Human ncp
3	1806	100.0	339	5	ABJ04073	Abj04073 Human G p
4	1806	100.0	339	5	ABG93791	Abg93791 Human G p
5	1806	100.0	339	5	ADH48916	Adh48916 NOV84B pr
6	1806	100.0	339	5	ADH48908	Adh48908 NOV81B pr
7	1806	100.0	339	5	ADH48906	Adh48906 NOV81A pr
8	1806	100.0	339	6	ABP81732	Abp81732 Human tra
9	1806	100.0	339	8	AD029662	Ad029662 Human GPC
10	1806	100.0	339	8	AD037691	Ad037691 Human tra
11	1801	99.7	338	4	AAB82970	Aab82970 Human G p
12	1791	99.2	339	5	ABB80694	Abb80694 Human tra
13	1539	85.2	311	7	ADCI2792	Adci2792 Human GPC
14	1525	84.4	296	4	AAG80972	Aag80972 Human ncp
15	1525	84.4	296	5	ABG93790	Abg93790 Human G p
16	1403	77.7	332	4	AAB49231	Aab49231 Rat SNORF
17	1333	77.1	332	5	ABB80695	Abb80695 Rat trace
18	1351	74.8	332	4	AAB49234	Aab49234 Mouse SNO
19	1351	74.8	332	8	AD029663	Ad029663 Mouse GPC
20	1212	67.1	238	4	AAG80960	Aag80960 Human ncp
21	1212	67.1	238	5	ABG93778	Abg93778 Human G p
22	998	55.3	191	4	AAB49230	Aab49230 Human SNO
23	923.5	51.1	347	8	AD029145	Ad029145 Mouse nov
24	886.5	49.1	343	8	AD029419	Ad029419 Mouse GPC
25	848	47.0	351	5	ABB79888	Abb79888 Human SNO

26	848	47.0	351	7	ADE81115	Ad81115 Human SNO
27	840.5	46.5	306	8	AD029074	Ad029074 Mouse nov
28	820	45.4	319	4	AAG72332	Aag72332 Human OR-
29	820	45.4	544	7	ADF70407	Adf70407 Orphan re
30	817	45.2	306	2	AAW55831	Aaw55831 Human G p
31	817	45.2	306	6	ABB58746	Abb58746 Human G-p
32	815	45.1	306	6	ABP57089	Abp57089 Human GPC
33	815	45.1	306	6	ABP81701	Abp81701 Human G p
34	815	45.1	306	8	AD005722	Ad005722 Human G p
35	815	45.1	306	8	AD029073	Ad029073 Human nov
36	814.5	45.1	295	4	AAG72401	Aag72401 Human OR-
37	813.5	45.0	343	2	AAW59907	Aaw59907 Human HNH
38	813.5	45.0	343	6	AAO19904	Aao19904 Human TA4
39	813.5	45.0	343	6	ABP81702	Abp81702 Human G p
40	813.5	45.0	343	8	AD029418	Ad029418 Human GPC
41	782	43.3	338	3	AAB18765	Aab18765 Amino aci
42	770	42.6	348	8	AD029148	Ad029148 Mouse nov
43	743	41.1	339	6	ABJ37891	Abj37891 NOVX prot
44	743	41.1	348	3	AAB18764	Aab18764 Amino aci
45	743	41.1	348	5	ABB98168	Abb98168 Human AXO

ALIGNMENTS

RESULT 1
ID AAB49232 standard; protein; 339 AA.
AC AAB49232;
XX
XX
DT 14-MAR-2001 (first entry)
XX
XX
DE Human SNORF33 receptor protein.
KW SNORF33; inflammation; arthritis; neurological disorder; infection;
KW bone disease; respiratory disorder; asthma; cancer; cardiovascular.
OS Homo sapiens.
XX
XX
PN WO200073449-A1.
XX
XX
PD 07-DEC-2000.
XX
XX
PF 26-MAY-2000; 2000WO-US014654.
XX
XX
PR 28-MAY-1999; 99US-00322257.
XX
XX
PR 06-OCT-1999; 99US-00413433.
XX
XX
PA (SYNA-) SYNAPTIC PHARM CORP.
XX
XX
PI Borowsky BE, Ogozalek KL, Jones KA;
XX
XX
DR WPI; 2001-025252/03.
XX
XX
PT Nucleic acid encoding a mammalian (human, rat and mouse) SNORF33 receptor
PT which is useful for designing drugs for treating conditions such as a
PT chronic and acute inflammation, arthritis, neurological disorders and
PT microbial infections.
XX
XX
PS Claim 9; Fig 6; 227p; English.
XX
XX
CC The present invention relates to a mammalian SNORF33 receptor. SNORF33
CC antagonists and agonists are used to treat abnormalities brought about by
CC increased or decreased activity of the mammalian SNORF33 receptor. The
CC receptor is useful as a tool for designing drugs for treating conditions
CC such as a chronic and acute inflammation, arthritis, neurological
CC disorders, microbial infections, bone diseases, respiratory disorders
CC such as asthma, cancers, cardiovascular disorders
XX
XX
SQ Sequence 339 AA:
Query Match 100.0%; Score 1806; DB 4; Length 339;

Best Local Similarity 100.0%; Pred. No. 1,4e-196; Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MMEPCNNINISCVKNNMSNDVRASLYSLMWLIIITLVGNLIYVISIHFQKQHTPTNW 60
Db 1 MMEPCNNINISCVKNNMSNDVRASLYSLMWLIIITLVGNLIYVISIHFQKQHTPTNW 60
QY 61 LIHSMATVDFLLGCLVMPYSMWRSAEHCWYFGEVFCIKHTSDIMLSSASIFHLSPISID 120
Db 61 LIHSMATVDFLLGCLVMPYSMWRSAEHCWYFGEVFCIKHTSDIMLSSASIFHLSPISID 120
QY 121 RYVAVCDPLRYKAKNNILVICWIFISWSPAVAFGMIFLELNFGKAEIYYKHYHCRG 180
Db 121 RYVAVCDPLRYKAKNNILVICWIFISWSPAVAFGMIFLELNFGKAEIYYKHYHCRG 180
QY 181 GCSVFESKISGVLTFMTSFYIPGSIIMLCVYRIYLIANEQARLISDANOKLQIGLEMXNG 240
Db 181 GCSVFESKISGVLTFMTSFYIPGSIIMLCVYRIYLIANEQARLISDANOKLQIGLEMXNG 240
QY 241 ISQSKERKAVKTLGIWGVFLICWCPFFICTVMDPFLHYIIPPTLNDVLIWFGYLNSTFN 300
Db 241 ISQSKERKAVKTLGIWGVFLICWCPFFICTVMDPFLHYIIPPTLNDVLIWFGYLNSTFN 300
QY 301 PMVYAFYFPMFRKALKMFLGKIFQKDSRCKLFLLELSS 339
Db 301 PMVYAFYFPMFRKALKMFLGKIFQKDSRCKLFLLELSS 339

RESULT 2
AAG80973 standard; protein; 339 AA.
XX
AC AAG80973;
XX
DT 28-AUG-2001 (first entry)
XX
DE Human ngPCR56 #3.
XX
KW G protein-coupled receptor; ngPCR; seven transmembrane receptor;
KW signal transduction; schizophrenia; thyroid disorder; renal failure;
KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;
KW cardiovascular disease; proliferative disorder; hormonal disorder;
KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
KW attention deficit-hyperactivity disorder; attention deficit disorder;
KW Parkinson's disease; migraine; senile dementia; inflammatory disease;
KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;
KW neuroprotective.
XX
OS Homo sapiens.
XX
PN WO200136473-A2.
XX
PD 25-MAY-2001.
XX
PF 16-NOV-2000; 2000WO-US031581.
XX
PR 16-NOV-1999; 99US-0165838P.
PR 17-NOV-1999; 99US-0166071P.
PR 19-NOV-1999; 99US-0166678P.
PR 28-DEC-1999; 99US-0173396P.
PR 22-FEB-2000; 2000US-0184129P.
PR 28-FEB-2000; 2000US-0185421P.
PR 28-FEB-2000; 2000US-0185549P.
PR 02-MAR-2000; 2000US-0186530P.
PR 03-MAR-2000; 2000US-0186811P.
PR 09-MAR-2000; 2000US-0188114P.
PR 17-MAR-2000; 2000US-0190310P.
PR 21-MAR-2000; 2000US-0190800P.
PR 20-APR-2000; 2000US-0198568P.
PR 02-MAY-2000; 2000US-0201190P.
PR 08-MAY-2000; 2000US-0203111P.
PR 25-MAY-2000; 2000US-0207094P.
XX

```

PA (PAA) PHARMACIA & UPJOHN CO.

XX Vogel I G, Wood LS, Parodi LA, Hiesch RR, Lind P, Slightow J,
 XX Schellin KA, Kayes PS, Bannigan CM, Ruff V, Seijtitz T, Huff RM,
 DR MPI: 2001-389826/41.
 XX N-PSDB; AAH51013.

PT New G protein-coupled receptor (ngPCR-x) and its encoding polynucleotide
 useful for diagnosing and treating e.g. schizophrenia.

XX Claim 37; Page 92; 261pp; English.

CC The present invention relates to novel G protein-coupled receptors
 CC (ngPCR); where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27, 28,
 CC 31-38, 40, 41, 53-60) and their coding sequences. The present sequence is
 CC one such G protein-coupled receptor. GPCRs are also known as seven
 CC transmembrane receptors and function in signal transduction. The ngPCR
 CC coding sequences are useful for screening a human to diagnose a disorder
 CC affecting the brain or a genetic predisposition, specifically
 CC schizophrenia. ngPCR are useful for identifying compounds useful for
 CC treating schizophrenia. Detection of ngPCR in a sample is useful as a
 CC diagnostic tool for diseases or disorders e.g. thyroid disorders, renal
 CC failure, rheumatoid arthritis, CNS disorders, infections such as HIV-1,
 CC metabolic and cardiovascular diseases, proliferative disorders and
 CC hormonal disorders. Modulators of ngPCR activity have the utility for
 CC treating neurological disorders, including schizophrenia, ADHD/ADD
 CC (attention deficit-hyperactivity disorder/attention deficit disorder),
 CC and neuronal disorders such as Alzheimer's disease, Parkinson's disease,
 CC migraine and senile dementia. Additional disorders include inflammatory
 CC conditions (e.g. Crohn's disease), rheumatoid arthritis, autoimmune
 CC disorders, cancers, respiratory ailments such as asthma, and inflammatory
 CC diseases e.g. inflammatory bowel disease

XX Sequence 339 AA;

Query Match 100.0%; Score 1806; DB 4; Length 339;

Best Local Similarity 100.0%; Pred. No. 1,4e-196; Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MMEPCNNINISCVKNNMSNDVRASLYSLMWLIIITLVGNLIYVISIHFQKQHTPTNW 60
Db 1 MMEPCNNINISCVKNNMSNDVRASLYSLMWLIIITLVGNLIYVISIHFQKQHTPTNW 60
QY 61 LIHSMATVDFLLGCLVMPYSMWRSAEHCWYFGEVFCIKHTSDIMLSSASIFHLSPISID 120
Db 61 LIHSMATVDFLLGCLVMPYSMWRSAEHCWYFGEVFCIKHTSDIMLSSASIFHLSPISID 120
QY 121 RYVAVCDPLRYKAKNNILVICWIFISWSPAVAFGMIFLELNFGKAEIYYKHYHCRG 180
Db 121 RYVAVCDPLRYKAKNNILVICWIFISWSPAVAFGMIFLELNFGKAEIYYKHYHCRG 180
QY 181 GCSVFESKISGVLTFMTSFYIPGSIIMLCVYRIYLIANEQARLISDANOKLQIGLEMXNG 240
Db 181 GCSVFESKISGVLTFMTSFYIPGSIIMLCVYRIYLIANEQARLISDANOKLQIGLEMXNG 240
QY 241 ISQSKERKAVKTLGIWGVFLICWCPFFICTVMDPFLHYIIPPTLNDVLIWFGYLNSTFN 300
Db 241 ISQSKERKAVKTLGIWGVFLICWCPFFICTVMDPFLHYIIPPTLNDVLIWFGYLNSTFN 300
QY 301 PMVYAFYFPMFRKALKMFLGKIFQKDSRCKLFLLELSS 339
Db 301 PMVYAFYFPMFRKALKMFLGKIFQKDSRCKLFLLELSS 339

RESULT 3
ABJ04073 standard; protein; 339 AA.
XX
AC ABJ04073;
XX
DT 11-OCT-2002 (first entry)
XX

```

DE Human G protein coupled receptor hRUP33.
XX Human; G-protein coupled receptor; GPCR; hRUP28; hRUP29; hRUP30; hRUP31;
KW hRUP32; hRUP33; hRUP34; hRUP35; hRUP36; hRUP37.
XX Homo sapiens.
PN WO200242461-A2.
XX
PD 30-MAY-2002.
XX
PF 26-NOV-2001; 2001WO-US044386.
XX
PR 27-NOV-2000; 2000US-0253404P.
PR 12-DEC-2000; 2000US-0255366P.
PR 20-FEB-2001; 2001US-0270266P.
PR 20-FEB-2001; 2001US-0270286P.
PR 06-APR-2001; 2001US-0282032P.
PR 06-APR-2001; 2001US-0282356P.
PR 06-APR-2001; 2001US-0282358P.
PR 06-APR-2001; 2001US-0282365P.
PR 14-MAY-2001; 2001US-0290917P.
PR 31-JUL-2001; 2001US-0309208P.
XX
XX (AREN-) ARENA PHARM INC.
XX
XX Chen R, Chu ZL, Dang HT, Lowitz KP, Pride C;
PI WPI; 2002-566565/60.
DR N-PSDB; ABT04871.
XX
XX Novel endogenous and non-endogenous versions of G protein-coupled
PT receptor useful for identification of candidate compounds as receptor
PT agonists or antagonists for use as therapeutic agents.
XX
XX Claim 21; Page 66-67; 84pp; English.
XX
XX The present invention provides the protein and coding sequences of
CC several human G-protein coupled receptors (GPCRs). These can be used in
CC the identification of candidate compounds as receptor agonists or inverse
CC agonists having applicability as therapeutic agents. The present sequence
CC is a GPCR protein of the invention
XX
XX Sequence 339 AA;
SQ
Query Match 100.0%; Score 1806; DB 5; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.4e-196;
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHPFGHNINISCVKNNMNDVRASLYSLMVLIIITLVGNLIVIVSISHPKQLHTPTM 60
DB 1 MHPFGHNINISCVKNNMNDVRASLYSLMVLIIITLVGNLIVIVSISHPKQLHTPTM 60
QY 61 LIHSMAVDVFLGCLVMPYSWVSAHCHWYFGEVFCIKTSTIDIMLSASIFHLSPISID 120
DB 61 LIHSMAVDVFLGCLVMPYSWVSAHCHWYFGEVFCIKTSTIDIMLSASIFHLSPISID 120
QY 121 RRYAACDPLRYAKAKNIIIVICWIFISNSVPAVFAAGMIFLEINFGAEEIYKHYHCHG 180
DB 121 RRYAACDPLRYAKAKNIIIVICWIFISNSVPAVFAAGMIFLEINFGAEEIYKHYHCHG 180
QY 181 GGSVFPFSKISGVLTFMTSTFYIPGSIWLCVYRIRYLAKQOARLISANOKLQIGLEMKNG 240
DB 181 GGSVFPFSKISGVLTFMTSTFYIPGSIWLCVYRIRYLAKQOARLISANOKLQIGLEMKNG 240
QY 241 ISQSKERKAVKTLGIWGVFLICWCPFFICTVMDPLAHYIIPPTLNDVLIWFGYLNSTN 300
DB 241 ISQSKERKAVKTLGIWGVFLICWCPFFICTVMDPLAHYIIPPTLNDVLIWFGYLNSTN 300
QY 301 PMVYAFYPMFRKALKQMLFGKIFQKSSRCKLFEILSS 339
DB 301 PMVYAFYPMFRKALKQMLFGKIFQKSSRCKLFEILSS 339

RESULT 4
ID ABG93791 standard; protein; 339 AA.
XX
XX ABG93791;
XX
DT 26-NOV-2002 (first entry)
XX
DE Human G protein-coupled receptor protein, nGPCR-56, #2.
XX
XX Human; receptor; G protein-coupled receptor; GPCR; nGPCR; beGPCR;
KW nG protein coupled receptor; communication; serpentine structure;
KW seven transmembrane receptor; 7TM; mental disorder; diagnosis;
KW genetic predisposition; brain; immune response; gene therapy;
KW anxiety disorder; depression; bipolar disorder; schizophrenia;
KW Huntington's disease; dyskinesia; manic depression; stroke;
KW Parkinson's disease; Alzheimer's disease; diabetes; inflammation; wound;
KW tranquilliser.
XX
XX Homo sapiens.
XX
XX WO200264789-A1.
XX
PD 22-AUG-2002. ✓
XX
PF 14-FEB-2001; 2001WO-US004641.
XX
PR 14-FEB-2001; 2001WO-US004641.
XX
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
XX Lind P, Parodi LA, Vogel G, Wood LS;
PI WPI; 2002-674879/72.
DR N-PSDB; ABS70246.
XX
XX New nucleic acids and polypeptides of the nG protein-coupled receptor,
PT useful for treating or diagnosing a mental disorder or a disorder
PT affecting the brain, e.g. anxiety disorders, schizophrenia, stroke or
PT Parkinson's disease.
XX
XX Example 1; Page 87; 244pp; English.
XX
XX The invention discloses an isolated human polypeptide, and encoding
CC nucleic acid, for a G protein-coupled receptor (GPCR), particularly the
CC nG protein coupled receptor-14 (nGPCR-14). GPCRs are vital in the
CC communication between cells and their environment and are characterised
CC by a serpentine structure that passes through the cell membrane seven
CC times, hence the reason such receptors are sometimes called seven
CC transmembrane receptors (7TM). The polynucleotides and polypeptides are
CC useful for identifying an nGPCR allelic variant that correlates with a
CC mental disorder, for isolating an antibody that binds to an epitope of
CC the polypeptide, for identifying a compound that binds the polypeptide or
CC polynucleotide and/or modulates its biological activity, for screening a
CC human subject to diagnose a disorder, or a genetic predisposition to a
CC disorder, affecting the brain or a genetic disposition to the disorder,
CC for identifying compounds useful for the treatment of a mental disorder
CC and for identifying a compound useful as a modulator of binding between
CC nGPCR-14 and a binding partner of nGPCR-14. The polypeptide is also
CC useful for inducing an immune response in a mammal. The nucleic acid or
CC polypeptide is particularly useful, using gene therapy, for treating e.g.
CC anxiety disorders, depression, bipolar disorder, schizophrenia,
CC Huntington's disease, dyskinesias, manic depression, stroke, Parkinson's
CC disease or Alzheimer's disease. The nucleic acid and polypeptide may also
CC be used for treating diabetes, inflammation or wounds. The sequences
CC presented in as beGPCRs) ABG93793, ABG93795 and ABG93796 are the nGPCR (also
XX referred to as beGPCRs) proteins
XX
XX Sequence 339 AA;
SQ
Query Match 100.0%; Score 1806; DB 5; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.4e-196;

Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPEPCNNINISCVKNNMSNDVRASLSLWVLIITLTGNTLVIVYISHPKQLHPTNW 60

Db 1 MPEPCNNINISCVKNNMSNDVRASLSLWVLIITLTGNTLVIVYISHPKQLHPTNW 60

QY 61 LIHSMATVDLGLGLVMPYSNVRSAEHCWYFGEVCKIHTSDIMLSASIFHLSFISID 120

Db 61 LIHSMATVDLGLGLVMPYSNVRSAEHCWYFGEVCKIHTSDIMLSASIFHLSFISID 120

QY 121 RYAVACDPLRYKAKMNLIVICWIFISWSVPAVAFGMIFLEINFGAEEIYKHHVCHRG 180

Db 121 RYAVACDPLRYKAKMNLIVICWIFISWSVPAVAFGMIFLEINFGAEEIYKHHVCHRG 180

QY 181 GCSVFSKISGVLTFMTSFYIPGSIMLCVYRIRYLLAKQARLISDANKLOIGLEMKNG 240

Db 181 GCSVFSKISGVLTFMTSFYIPGSIMLCVYRIRYLLAKQARLISDANKLOIGLEMKNG 240

QY 241 ISQSKERKAVKTLGIWGVFLICWCPFFICTVNDPFLHYIIPPTLNDVLIWFGYLNSTN 300

Db 241 ISQSKERKAVKTLGIWGVFLICWCPFFICTVNDPFLHYIIPPTLNDVLIWFGYLNSTN 300

QY 301 PMVYAFYFPMFRKALKMFLGKIFQKDSRCKLFLFLSS 339

Db 301 PMVYAFYFPMFRKALKMFLGKIFQKDSRCKLFLFLSS 339

RESULT 5
ADH48916
ID ADH48916 standard; protein; 339 AA.
XX
AC ADH48916;
XX
DT 25-MAR-2004 (first entry)
XX
DE NOV84B protein sequence, SEQ ID 200.
XX
KW Human; NOVX; atherosclerosis; hypertension; obesity; cancer; cytostatic;
KW hypotensive; antiarteriosclerotic; anorectic; gene therapy; NOV84B;
KW GPCR-like protein; chromosome 6.
XX
OS Homo sapiens.
XX
PN WO200268652-A2.
XX
PD 06-SEP-2002. ✓
XX
PF 26-FEB-2002; 2002WO-US005910.
XX
PR 26-FEB-2001; 2001US-0271646P.
PR 27-FEB-2001; 2001US-0271840P.
PR 28-FEB-2001; 2001US-0272404P.
PR 28-FEB-2001; 2001US-0272405P.
PR 28-FEB-2001; 2001US-0272410P.
PR 28-FEB-2001; 2001US-0272414P.
PR 02-MAR-2001; 2001US-0272787P.
PR 02-MAR-2001; 2001US-0272922P.
PR 02-MAR-2001; 2001US-0273048P.
PR 02-MAR-2001; 2001US-0273300P.
PR 16-MAR-2001; 2001US-0276401P.
PR 20-MAR-2001; 2001US-0278660P.
PR 30-MAR-2001; 2001US-0280039P.
PR 30-MAR-2001; 2001US-0280234P.
PR 02-APR-2001; 2001US-0280818P.
PR 12-APR-2001; 2001US-0283443P.
PR 23-APR-2001; 2001US-0285754P.
PR 24-APR-2001; 2001US-0286096P.
PR 03-MAY-2001; 2001US-0288353P.
PR 17-MAY-2001; 2001US-0291703P.
PR 31-MAY-2001; 2001US-0294834P.
PR 20-JUN-2001; 2001US-0296959P.
PR 21-JUN-2001; 2001US-0299845P.

PR 05-JUL-2001; 2001US-0303242P.
PR 13-AUG-2001; 2001US-0311981P.
PR 16-AUG-2001; 2001US-0312858P.
PR 17-AUG-2001; 2001US-0313280P.
PR 29-AUG-2001; 2001US-0315614P.
PR 17-SEP-2001; 2001US-0322818P.
PR 25-FEB-2002; 2002US-00322818.
XX
XX (CURA-) CUBAGEN CORP.
XX
PI AIsbrook JP, Anderson DW, Ballinger RA, Boldog FL, Burgess CE;
PI Casman SJ, Sillerman KE, Gangoli BA, Gerlach VL, Gilbert JA;
PI Gorman L, Guo X, Guzey VY, Kekuda R, Li L, Liu X, Malpankar UM;
PI Miller CE, Millet I, Padigaru M, Patcurajan M, Pena CE, Peyman JA;
PI Rastelli L, Shenoy SG, Shimkets RA, Smithson G, Szytek KA, Stone DJ;
PI Taupier RJ, Tchernev VT, Vernet CAM, Zernušen BD;
XX
DR WPI; 2002-698672/75.
DR N-PsDB; ADH48915.
XX
XX
PT New NOVX polypeptides or polynucleotides, useful for preventing or
PT treating disorders or syndromes e.g., atherosclerosis, hypertension,
PT obesity or cancer.
XX
PS Claim 1; Page 452; 923pp; English.
XX
XX
CC The present invention relates to novel human NOVX proteins, where X is
CC any number from 1 to 91 and their coding sequences. The proteins and
CC coding sequences are useful for preventing or treating disorders or
CC syndromes e.g., atherosclerosis, hypertension, obesity or cancer. NOVX3-
CC NOV91 are GPCR-like proteins. The coding sequence for the present NOV
CC maps to chromosome 6.
XX
SQ Sequence 339 AA:

Query Match 100.0%; Score 1806; DB 5; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.4e-196;
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPEPCNNINISCVKNNMSNDVRASLSLWVLIITLTGNTLVIVYISHPKQLHPTNW 60

Db 1 MPEPCNNINISCVKNNMSNDVRASLSLWVLIITLTGNTLVIVYISHPKQLHPTNW 60

QY 61 LIHSMATVDLGLGLVMPYSNVRSAEHCWYFGEVCKIHTSDIMLSASIFHLSFISID 120

Db 61 LIHSMATVDLGLGLVMPYSNVRSAEHCWYFGEVCKIHTSDIMLSASIFHLSFISID 120

QY 121 RYAVACDPLRYKAKMNLIVICWIFISWSVPAVAFGMIFLEINFGAEEIYKHHVCHRG 180

Db 121 RYAVACDPLRYKAKMNLIVICWIFISWSVPAVAFGMIFLEINFGAEEIYKHHVCHRG 180

QY 181 GCSVFSKISGVLTFMTSFYIPGSIMLCVYRIRYLLAKQARLISDANKLOIGLEMKNG 240

Db 181 GCSVFSKISGVLTFMTSFYIPGSIMLCVYRIRYLLAKQARLISDANKLOIGLEMKNG 240

QY 241 ISQSKERKAVKTLGIWGVFLICWCPFFICTVNDPFLHYIIPPTLNDVLIWFGYLNSTN 300

Db 241 ISQSKERKAVKTLGIWGVFLICWCPFFICTVNDPFLHYIIPPTLNDVLIWFGYLNSTN 300

QY 301 PMVYAFYFPMFRKALKMFLGKIFQKDSRCKLFLFLSS 339

Db 301 PMVYAFYFPMFRKALKMFLGKIFQKDSRCKLFLFLSS 339

RESULT 6
ADH48908
ID ADH48908 standard; protein; 339 AA.
XX
AC ADH48908;
XX
DT 25-MAR-2004 (first entry)
XX
DE NOV81B protein sequence, SEQ ID 192.

KM	Human, NOVX; atherosclerosis; hypertension; obesity; cancer; cytostatic;
KW	hypotensive; antiplateletisoleptic; anorectic; gene therapy; NOV81B.
KM	GPCR-like protein; chromosome 6.
XX	
OS	Homo sapiens.
PN	W0200268652-A2.
XX	/
PD	06-SEP-2002.
XX	
PF	26-FEB-2002; 2002MO-US050910.
XX	
PR	26-FEB-2001; 2001US-0271646P.
PR	27-FEB-2001; 2001US-0271640P.
PR	28-FEB-2001; 2001US-0272405P.
PR	28-FEB-2001; 2001US-0272405P.
PR	28-FEB-2001; 2001US-0272410P.
PR	28-FEB-2001; 2001US-0272410P.
PR	02-MAR-2001; 2001US-0272787P.
PR	02-MAR-2001; 2001US-0272922P.
PR	02-MAR-2001; 2001US-0273048P.
PR	02-MAR-2001; 2001US-0273300P.
PR	16-MAR-2001; 2001US-0276401P.
PR	20-MAR-2001; 2001US-0277324P.
PR	20-MAR-2001; 2001US-0278660P.
PR	30-MAR-2001; 2001US-0280039P.
PR	30-MAR-2001; 2001US-0280234P.
PR	02-APR-2001; 2001US-0280818P.
PR	12-APR-2001; 2001US-0283443P.
PR	23-APR-2001; 2001US-0285754P.
PR	24-APR-2001; 2001US-0286096P.
PR	03-MAY-2001; 2001US-0288353P.
PR	17-MAY-2001; 2001US-0291703P.
PR	31-MAY-2001; 2001US-0294834P.
PR	20-JUN-2001; 2001US-0296695P.
PR	21-JUN-2001; 2001US-0299845P.
PR	05-JUL-2001; 2001US-0303242P.
PR	13-AUG-2001; 2001US-0311981P.
PR	16-AUG-2001; 2001US-0312858P.
PR	17-AUG-2001; 2001US-0313280P.
PR	29-AUG-2001; 2001US-0315614P.
PR	17-SEP-2001; 2001US-0322818P.
PR	25-FEB-2002; 2002US-00322818.
XX	
PA	(CURA-) CURAGEN CORP.
PI	A1sobrook JP, Anderson DW, Ballinger RA, Boldog FL, Burgess CE;
PI	Cabman SJ, Ellerman KE, Gangoli EA, Gerlach VL, Gilbert JA;
PI	Gorman SJ, Guo X, Guev PY, Kekuda R, Li L, Liu X, Malyankar UM;
PI	Miller CE, Millet I, Padigar M, Paturajan M, Pena CE, Peyman JA;
PI	Racelli L, Shenoy SG, Shinkens RA, Smithson G, Splet KA, Stone DJ,
PI	Taupier RJ, Tchernov VT, Vernet CM, Zernusen BD;
XX	
DR	WPI; 2002-698672/75.
DR	N-PSDB; ADH48907.
PT	New NOVX polypeptides or polymucleotides, useful for preventing or
PT	treating disorders or syndromes e.g., atherosclerosis, hypertension,
FT	obesity or cancer.
XX	
PS	Claim 1; Page 440; 923pp; English.
CC	The present invention relates to novel human NOVX proteins, where X is
CC	any number from 1 to 91 and their coding sequences. The proteins and
CC	coding sequences are useful for preventing or treating disorders or
CC	syndromes e.g., atherosclerosis, hypertension, obesity or cancer. NOV3-
CC	NOV91 are GPCR-like proteins. The coding sequence for the present NOV
CC	maps to chromosome 6.
XX	
SQ	Sequence 339 AA;

Query Match 100.0%; Score 1806; DB 5; Length 339.

	Best Local Similarity	100.0%	Pred. No. 1,4e-196		Matches 339	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1	MMPECHNIINISCVKNNMNSDVRASLSYSLMWLIILTTLVGNLIIVISISHFKQLHTPTNW	60						
Db	1	MMPECHNIINISCVKNNMNSDVRASLSYSLMWLIILTTLVGNLIIVISISHFKQLHTPTNW	60						
Qy	61	LIHSMATYDPLFLGCLWMPYSNVRASHEHCWYRGEVCKIHTSTDIIMLSASIFHLSPISID	120						
Db	61	LIHSMATYDPLFLGCLWMPYSNVRASHEHCWYRGEVCKIHTSTDIIMLSASIFHLSPISID	120						
Qy	121	RYVAVCDPLRYKAKNNILVICMFIPISSVAVFAFGMIFLELNFGAEEIYYKHHVCRG	180						
Db	121	RYVAVCDPLRYKAKNNILVICMFIPISSVAVFAFGMIFLELNFGAEEIYYKHHVCRG	180						
Qy	181	GCSEVFSKISGVLTFMTSFTYIPGSIIMLCVYRIYILAKEQARLISDANQKQIGLEMKRG	240						
Db	181	GCSEVFSKISGVLTFMTSFTYIPGSIIMLCVYRIYILAKEQARLISDANQKQIGLEMKRG	240						
Qy	241	ISQSKERAAVKTLLGIVMGVFLICWCPFFICTYMDPFLHYIIPPTLNDVLIWFGYLNSTN	300						
Db	241	ISQSKERAAVKTLLGIVMGVFLICWCPFFICTYMDPFLHYIIPPTLNDVLIWFGYLNSTN	300						
Qy	301	PMVYAFEPYPMWRKALKXMLFGKIPQKSSRCCLFLELSS	339						
Db	301	PMVYAFEPYPMWRKALKXMLFGKIPQKSSRCCLFLELSS	339						
RESULT 7									
ADH48906									
ID	ADH48906	standard; protein, 339 AA.							
XX	ADH48906;								
AC									
XX									
DT	25-MAR-2004	(first entry)							
XX									
DE	NOV81A	protein sequence, SEQ ID 190.							
XX									
KW	Human; NOXV; atherosclerosis; hypertension; obesity; cancer; cytostatic;								
KW	hypotensive; antiarteriosclerotic; anorectic; gene therapy; NOV81A;								
XX	GPCR-like protein; chromosome 6.								
XX									
OS	Homo sapiens.								
XX									
PN	MO200268652-A2.								
XX									
PD	06-SEP-2002.	/							
XX									
PF	26-FEB-2002;	2002WO-US0005910.							
XX									
FR	26-FEB-2001;	2001US-0271646P.							
FR	27-FEB-2001;	2001US-0271840P.							
PR	28-FEB-2001;	2001US-0272404P.							
PR	28-FEB-2001;	2001US-0272405P.							
PR	28-FEB-2001;	2001US-0272410P.							
PR	28-FEB-2001;	2001US-0272414P.							
PR	02-MAR-2001;	2001US-0272787P.							
PR	02-MAR-2001;	2001US-0272922P.							
PR	02-MAR-2001;	2001US-0273048P.							
PR	02-MAR-2001;	2001US-0273300P.							
PR	16-MAR-2001;	2001US-0276401P.							
PR	20-MAR-2001;	2001US-0277324P.							
PR	20-MAR-2001;	2001US-0278660P.							
PR	30-MAR-2001;	2001US-0280039P.							
PR	30-MAR-2001;	2001US-0280234P.							
PR	02-APR-2001;	2001US-0280818P.							
PR	12-APR-2001;	2001US-0283443P.							
PR	23-APR-2001;	2001US-0285754P.							
PR	24-APR-2001;	2001US-0286096P.							
PR	03-MAY-2001;	2001US-0286353P.							
PR	17-MAY-2001;	2001US-0291703P.							
PR	31-MAY-2001;	2001US-0296834P.							
PR	20-JUN-2001;	2001US-0299695P.							

```

PR 21-JUN-2001; 2001US-0299845P.
PR 05-JUL-2001; 2001US-0303242P.
PR 13-AUG-2001; 2001US-0311981P.
PR 16-AUG-2001; 2001US-0312858P.
PR 17-AUG-2001; 2001US-0313280P.
PR 29-AUG-2001; 2001US-0315614P.
PR 17-SEP-2001; 2001US-0322818P.
PR 25-FEB-2002; 2002US-00322818.
XX
XX (CURA-) CURAGEN CORP.
XX
PI Alsobrook JP, Anderson DW, Ballinger RA, Boldog FL, Burgess CE,
PI Gorman SO, Ellerman KE, Gangoli EA, Gerlach VL, Gilbert JA,
PI Gorman L, Guo X, Guev VY, Kikuda R, Li L, Liu X, Malysankar UM,
PI Miller CE, Millet I, Padigar M, Paturajan M, Pena CE, Peyman JA,
PI Rastelli L, Shenoy SG, Shinkens RA, Smithson G, Spytek KA, Stone DJ,
PI Taupier RJ, Tchernev VT, Vernet CM, Zernhusen BD;
XX
DR WPI; 2002-698672/75.
DR N-PSDB; ADH48905.
XX
PT New NOVX polypeptides or polynucleotides, useful for preventing or
PT treating disorders or syndromes e.g., atherosclerosis, hypertension,
PT obesity or cancer.
XX
PS Claim 1; Page 439; 923pp; English.
XX
CC The present invention relates to novel human NOVX proteins, where X is
CC any number from 1 to 91 and their coding sequences. The proteins and
CC coding sequences are useful for preventing or treating disorders or
CC syndromes e.g. atherosclerosis, hypertension, obesity or cancer. NOVX-
CC NOV91 are GPCR-like proteins. The coding sequence for the present NOV
CC maps to chromosome 6.
XX
SQ Sequence 339 AA;

```

```

Query Match          100.0%; Score 1806; DB 5; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.4e-196;
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MMPECHNINISCVKNNMSNDVRASLYSLMVLIIITLWGNLIIVYSISHFQQLHTPTNW 60
DB 1 MMPECHNINISCVKNNMSNDVRASLYSLMVLIIITLWGNLIIVYSISHFQQLHTPTNW 60
QY 61 LIHSAVYDFLLGLCLVMPYSWVSAEHQWYEGVEVCKIHTSTDIMLSASIFHLSFISID 120
DB 61 LIHSAVYDFLLGLCLVMPYSWVSAEHQWYEGVEVCKIHTSTDIMLSASIFHLSFISID 120
QY 121 RYVACDPLRYAKKNIIIVICMIFISVPAVFAFGMIFELNKGABEIIYKXVHCRG 180
DB 121 RYVACDPLRYAKKNIIIVICMIFISVPAVFAFGMIFELNKGABEIIYKXVHCRG 180
QY 121 RYVACDPLRYAKKNIIIVICMIFISVPAVFAFGMIFELNKGABEIIYKXVHCRG 180
DB 121 RYVACDPLRYAKKNIIIVICMIFISVPAVFAFGMIFELNKGABEIIYKXVHCRG 180
QY 181 GCSVFPKISGVLTFTMTSFYIPGSIIMLCVYRIYLIKEQARLISDAOKIQIGEMKNG 240
DB 181 GCSVFPKISGVLTFTMTSFYIPGSIIMLCVYRIYLIKEQARLISDAOKIQIGEMKNG 240
QY 241 ISQSKERAAVKTLLGIVMGVFLICWCPFFICTVMDPFLHYIIPPTLNDVLIWFGYLNSTFN 300
DB 241 ISQSKERAAVKTLLGIVMGVFLICWCPFFICTVMDPFLHYIIPPTLNDVLIWFGYLNSTFN 300
QY 301 PMVYAFYFPMFRKALIKMMLFKKIPKOKSSRCKLPLELSS 339
DB 301 PMVYAFYFPMFRKALIKMMLFKKIPKOKSSRCKLPLELSS 339

```

```

RESULT 8
ABP81732
ID ABP81732 standard; protein; 339 AA.
XX
AC ABP81732;
XX
DT 04-MAR-2003 (first entry)
XX

```

```

DE Human trace amine receptor 1 protein SEQ ID NO:639.
XX
XX G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
XX G protein-coupled receptor modulator; antibody; immune-related disease;
XX growth-related disease; cell regeneration-related disease; AIDS; cancer;
XX immunological-related disease; cell proliferative disease; autoimmune disease;
XX Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
XX osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
XX gastritis; anxiety; depression; Parkinson's disease; multiple sclerosis; pain;
XX mental retardation; epilepsy; asthma; tuberculosis; obesity; memory loss;
XX hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
XX ulcer.
XX
OS Homo sapiens.
XX
PN WO200261087-A2.
XX
PD 08-AUG-2002.
XX
PF 19-DEC-2001; 2001WO-US050107.
XX
PR 19-DEC-2000; 2000US-0257144P.
XX
PA (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI Burner GC, Roush CL, Brown JP;
XX
PI WPI; 2003-046718/04.
XX
DR N-PSDB; ABZ42578.
XX
PT New isolated antigenic peptides e.g., for G protein-coupled receptors
PT (GPCR), useful for diagnosing and designing drugs for treating conditions
PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
PT autoimmune diseases.
XX
PS Disclosure; Fig 1; 523pp; English.
XX
CC The present invention describes antigenic peptides (I) comprising: (a)
CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
CC acids. Also described: (1) an assay for the detection of a particular G
CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC and (2) an isolated antibody having high specificity and high affinity or
CC avidity for a particular GPCR. (I) can be used as GPCR modulators and in
CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
CC antibody against a particular GPCR, and in the production of specific
CC antibodies. The peptides and antibodies are also useful for detecting the
CC presence or absence of corresponding GPCRs. The antigenic peptides for
CC GPCRs and antibodies are useful for diagnosing and designing drugs for
CC treating immune-related diseases, growth-related diseases, cell
CC regeneration-related disease, immunological-related cell proliferative
CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC any other disorder in which GPCRs are involved. The antibodies may be
CC used in immunoassays and immunodiagnoses. ABZ42523 to ABZ42869 encode
CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC exemplification of the present invention
XX
SQ Sequence 339 AA;

```

```

Query Match          100.0%; Score 1806; DB 6; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.4e-196;
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MMPECHNINISCVKNNMSNDVRASLYSLMVLIIITLWGNLIIVYSISHFQQLHTPTNW 60
DB 1 MMPECHNINISCVKNNMSNDVRASLYSLMVLIIITLWGNLIIVYSISHFQQLHTPTNW 60

```

QY 61 LIHSMATVDFLLGLVMPYSAHCHWYFGEVFCCKIHTSTDIMSSASIFHLSFISID 120
 CC 61 LIHSMATVDFLLGLVMPYSAHCHWYFGEVFCCKIHTSTDIMSSASIFHLSFISID 120
 Db 61 LIHSMATVDFLLGLVMPYSAHCHWYFGEVFCCKIHTSTDIMSSASIFHLSFISID 120
 QY 121 RYAVACDPLRYAKANNIIVICWIFISWVPAVAFGMIFLELNFGAEIYYKKVHCRCG 180
 CC 121 RYAVACDPLRYAKANNIIVICWIFISWVPAVAFGMIFLELNFGAEIYYKKVHCRCG 180
 Db 121 RYAVACDPLRYAKANNIIVICWIFISWVPAVAFGMIFLELNFGAEIYYKKVHCRCG 180
 QY 181 GCSVFFSKISGVLTFMTSFYIPGSIMLCVYRYIYIAEQARLISDAQKQIGLEMKNG 240
 CC 181 GCSVFFSKISGVLTFMTSFYIPGSIMLCVYRYIYIAEQARLISDAQKQIGLEMKNG 240
 Db 181 GCSVFFSKISGVLTFMTSFYIPGSIMLCVYRYIYIAEQARLISDAQKQIGLEMKNG 240
 QY 241 ISOSKERAQVKTGLIGVGVFLICWCPFFICTYMDPFLHIIPPTLNDVLIWFGYLNSTFN 300
 CC 241 ISOSKERAQVKTGLIGVGVFLICWCPFFICTYMDPFLHIIPPTLNDVLIWFGYLNSTFN 300
 Db 241 ISOSKERAQVKTGLIGVGVFLICWCPFFICTYMDPFLHIIPPTLNDVLIWFGYLNSTFN 300
 QY 301 PMVYAFYPWPFRKALKMMLFGKIFOKDSRCKLPLELSS 339
 CC 301 PMVYAFYPWPFRKALKMMLFGKIFOKDSRCKLPLELSS 339
 Db 301 PMVYAFYPWPFRKALKMMLFGKIFOKDSRCKLPLELSS 339
 RESULT 9
 ID ADO29662 standard; protein; 339 AA.
 AC ADO29662;
 XX 29-JUL-2004 (first entry)
 DT Human GPCR TAA1, SEQ ID NO:764.
 DE
 XX G protein-coupled receptor; GPCR; drug screening; diagnosis;
 KW transgenic mouse; neurological disorder; adrenal gland disorder;
 KW colon disorder; intestinal disorder; cardiovascular disorder;
 KW muscular disorder; blood disorder; immune disorder; bone disorder;
 KW joint disorder; metabolic disorder; nutritive disorder; cancer;
 KW kidney disorder; liver disorder; lung disorder; breast disorder;
 KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
 KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
 KW thymus disorder; thyroid disorder; antiparkinsonian; anti-nausea;
 KW cytoskeletal; antiinflammatory; vasotropic; antiangiinal; antidiabetic;
 KW CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;
 KW vitruclide; hepatotropic; antibacterial; antianemic; antiseborrheic;
 KW dermatological; antidiarr; antithyroid; antiallergic; anorectic;
 KW immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
 KW receptor.
 KW Homo sapiens.
 OS
 XX WO2004040000-A2.
 PN 13-MAY-2004.
 PD 09-SEP-2003; 2003WO-US028226.
 PF 09-SEP-2002; 2003US-0409303P.
 PR 09-APR-2003; 2003US-0461329P.
 XX (PRIM-) PRIMAL INC.
 PA Galanarlis GA, Bergmann JE, Gragerov A, Hohmann J, Li F;
 PI Madisen L, Melinwein KL, Pavlova MN, Vasiliadis D, Zeng H;
 XX WPI, 2004-390329/36.
 DR N-PSDB; ADO30050.
 XX Novel mammalian G protein coupled receptors, useful for identifying
 PT compounds that modulates diagnosing and treating disease condition
 PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
 PT pectoris, Parkinson's disease.
 XX Claim 151; SEQ ID NO 764; 542bp; English.

XX The invention relates to human and mouse G protein-coupled receptors
 CC (GPCRs) and nucleic acids encoding them. The invention also relates to
 CC sequences at least 90% identical to the GPCR proteins and nucleic acids
 CC of the invention, methods of treating, preventing or diagnosing diseases
 CC associated with GPCRs of the invention, methods of screening for
 CC compounds useful in the treatment of GPCR-related diseases; a transgenic
 CC mouse comprising a GPCR gene of the invention; a mouse comprising a
 CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
 CC from the transgenic mice; kits comprising several mice, each of which has
 CC a mutation in a different GPCR gene of the invention; and kits comprising
 CC probes which hybridise to GPCR polynucleotides of the invention. The
 CC invention further discloses variants of the GPCR polypeptides and vectors
 CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
 CC be used in the diagnosis, treatment or prevention of a wide variety of
 CC diseases including neurological disorders (e.g., Alzheimer's disease,
 CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
 CC disorders of the adrenal gland; disorders of the colon or intestine
 CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
 CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
 CC myocardial infarction); muscular disorders; blood disorders (e.g.,
 CC anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
 CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
 CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
 CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
 CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
 CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
 CC thyroid (e.g., cancers). The present sequence represents a GPCR of the
 CC invention. Note: The full sequence data for this patent did not form part
 CC of the printed specification; those sequences not shown were obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pcr_sequences.
 XX
 SQ Sequence 339 AA;
 Query Match 100.0%; Score 1806; DB 8; Length 339;
 Best Local Similarity 100.0%; Pred. No. 1,4e-196;
 Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MMPFCHNINISCVKNNMSNDVRABLYSLMVLITTLVGNLIYVVISHPQLHTPTNW 60
 Db 1 MMPFCHNINISCVKNNMSNDVRABLYSLMVLITTLVGNLIYVVISHPQLHTPTNW 60
 QY 61 LIHSMATVDFLLGLVMPYSAHCHWYFGEVFCCKIHTSTDIMSSASIFHLSFISID 120
 Db 61 LIHSMATVDFLLGLVMPYSAHCHWYFGEVFCCKIHTSTDIMSSASIFHLSFISID 120
 QY 121 RYAVACDPLRYAKANNIIVICWIFISWVPAVAFGMIFLELNFGAEIYYKKVHCRCG 180
 Db 121 RYAVACDPLRYAKANNIIVICWIFISWVPAVAFGMIFLELNFGAEIYYKKVHCRCG 180
 QY 181 GCSVFFSKISGVLTFMTSFYIPGSIMLCVYRYIYIAEQARLISDAQKQIGLEMKNG 240
 Db 181 GCSVFFSKISGVLTFMTSFYIPGSIMLCVYRYIYIAEQARLISDAQKQIGLEMKNG 240
 QY 241 ISOSKERAQVKTGLIGVGVFLICWCPFFICTYMDPFLHIIPPTLNDVLIWFGYLNSTFN 300
 Db 241 ISOSKERAQVKTGLIGVGVFLICWCPFFICTYMDPFLHIIPPTLNDVLIWFGYLNSTFN 300
 QY 301 PMVYAFYPWPFRKALKMMLFGKIFOKDSRCKLPLELSS 339
 Db 301 PMVYAFYPWPFRKALKMMLFGKIFOKDSRCKLPLELSS 339
 RESULT 10
 ID ADS73691 standard; protein; 339 AA.
 AC ADS73691;
 XX 16-DEC-2004 (first entry)
 DT Human trace amine receptor 1 (TA1) polypeptide.
 XX

XX TAL; GPCR; trace amine receptor 1; antianemic; antiinflammatory;
KW cardiovascular; neuroprotective; respiratory; uropathic;
KW G protein coupled receptor; gene therapy; human; CSRL.
OS Homo sapiens.
PN MO2004083851-A2.
XX
PD 30-SEP-2004.
XX
PF 04-MAR-2004; 2004MO-EP002182.
XX
PR 18-MAR-2003; 2003EP-00006101.
XX
PA (FARB) BAYER HEALTHCARE AG.
XX
PI Golz S, Brueggemeier U, Geerts A;
XX
DR WPI; 2004-691179/67.
XX
DR N-PSDB; ADS73690.
XX
PT Screening for therapeutic agents, useful for treating e.g.,
PT cardiovascular and inflammatory disorders, comprises contacting a test
PT compound with trace amine receptor 1 (TAR) polypeptide and detecting
PT their binding.
XX
PS Disclosure; SEQ ID NO 2; 117P; English.
XX
CC The invention relates to screening for therapeutic agents and involves
CC contacting a test compound with trace amine receptor 1 (TAR) polypeptide
CC or polynucleotide and detecting binding of the test compound to TAR
CC polypeptide or polynucleotide, or determining TAR polypeptide activity at
CC a certain test compound concentration or in the absence of the test
CC compound and at a different concentration of the test compound. The
CC therapeutic agents are useful in treating disease such as cardiovascular
CC disorders, neurological disorders, hematological disorders, respiratory
CC disorders, neurological disorders or urological disorders in a mammal. The
CC regulators of TAR are useful in the preparation of a pharmaceutical
CC composition for treating the above diseases and also useful for the
CC regulation of TAR activity in a mammal having the disease. The nucleotide
CC sequences encoding TAR are useful as hybridization probes, in
CC constructing oligomers for PCR, for chromosome and gene mapping, in the
CC recombinant production of TAR, in generating antisense DNA or RNA and in
CC molecular biology techniques that have not yet been developed. TAR are
CC useful for immunizing a mammal to produce polyclonal antibodies and for
CC diagnostic purposes. The present sequence represents a human TAR
CC receptor, a seven transmembrane G protein coupled receptor polypeptide.
XX
SQ Sequence 339 AA;
Query Match 100.0%; Score 1806; DB 8; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.4e-196;
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 241 ISQSKERKAVKTLGIWGVFLICWCPFFICTVMDPFLHYIIPTLNDVLIWGYLNSTFN 300
QY 301 PMVYAFPPWPRKALKMFLFGKIFQKDSRCKLFIELSS 339
DB 301 PMVYAFPPWPRKALKMFLFGKIFQKDSRCKLFIELSS 339
RESULT 11
AAB82970
ID AAB82970 standard; protein, 338 AA.
XX
XX AAB82970;
XX
XX 21-DEC-2001 (first entry)
XX
XX Human G protein coupled receptor.
XX
XX G protein coupled receptor; GPCR; aminergic receptor;
KW signal transduction; therapy; human.
XX
OS Homo sapiens.
FH Key
FT Modified-site
FT /note= "Asn is N-glycosylated"
FT 16..19
FT Modified-site
FT /note= "Asn is N-glycosylated"
FT 27..47
FT Domain
FT /note= "transmembrane domain 1"
FT 61..81
FT /note= "transmembrane domain 2"
FT 82..89
FT /note= "Tyr is O-phosphorylated by tyrosine kinase"
FT 99..119
FT /note= "transmembrane domain 3"
FT 99..102
FT /note= "Thr is O-phosphorylated by casein kinase II"
FT 108..124
FT Peptide
FT /note= "G protein coupled receptor signature"
FT 136..156
FT Domain
FT /note= "transmembrane domain 4"
FT 165..172
FT /note= "Tyr is O-phosphorylated by tyrosine kinase"
FT 189..209
FT Domain
FT /note= "transmembrane domain 5"
FT 239..244
FT Modified-site
FT /note= "N-myristoylated"
FT 252..272
FT Domain
FT /note= "transmembrane domain 6"
FT 253..258
FT Modified-site
FT /note= "N-myristoylated"
FT 270..273
FT Modified-site
FT /note= "Thr is O-phosphorylated by casein kinase II"
FT 283..286
FT Modified-site
FT /note= "Thr is O-phosphorylated by casein kinase II"
FT 287..307
FT Modified-site
FT /note= "transmembrane domain 7"
FT 295..298
FT Modified-site
FT /note= "Asn is N-glycosylated"
FT 327..329
FT /note= "Ser is O-phosphorylated by protein kinase C"
PN MO200172841-A2.
XX
XX 04-OCT-2001.
XX
XX 27-MAR-2001; 2001MO-US009660.
XX
XX 27-MAR-2000; 2000US-0192311P.
PR 04-AUG-2000; 2000US-00633145.
XX
XX (PEKE) PE CORP NY.
PA

PI Kodira C, Cravchik A, Di Francesco V, Beasley EM;
 XX
 DR WPI; 2001-611617/70.
 DR N-PSDB; AAH26919, AAH26919.
 XX
 PT New human G-protein coupled receptor, useful for identifying specific
 PT modulators, potential therapeutic agents, is related to the aminergic
 PT receptor family.
 XX
 PS Claim 1; Fig 2; 60pp; English.
 XX
 CC The present sequence is that of a novel human G protein coupled receptor
 CC (GPCR) that is related to aminergic receptor family. GPCRs, particularly
 CC members of the aminergic receptor family, are a major target for drug
 CC action and development. The GPCR polypeptides and polynucleotides of the
 CC invention are therefore useful for the development of human therapeutic
 CC targets, can aid in the identification of therapeutic proteins, and serve
 CC as targets for the development of human therapeutic agents. The present
 CC GPCR is expressed in foetal brain, brain, placenta, liver, stomach and
 CC kidney, and is involved in signal transduction pathways. By targeting an
 CC agent to modulate the GPCR, the signalling activity and biological
 CC process mediated by the receptor can be agonised or antagonised in
 CC specific cells and tissues in order to modulate its biological activity
 CC in a therapeutic context (mammalian therapy) or toxic context (anti-cell
 CC therapy e.g. anti-cancer therapy)
 CC
 SQ Sequence 338 AA;
 XX
 Query Match 99.7%; Score 1801; DB 4; Length 338;
 Best Local Similarity 100.0%; Pred. No. 5,1e-196;
 Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 2 MPFCNININISCVKNNMSNDVRASLSLMVLIIITLVGNLIVVISISHFQKQHTPTNW 61
 Db 1 MPFCNININISCVKNNMSNDVRASLSLMVLIIITLVGNLIVVISISHFQKQHTPTNW 60
 Oy 62 IHSMAVDPLGCLVMPYSWVRSAEHCWYGEVFCIKHTSTIDIMLSASIFHLSFISIDR 121
 Db 61 IHSMAVDPLGCLVMPYSWVRSAEHCWYGEVFCIKHTSTIDIMLSASIFHLSFISIDR 120
 Oy 122 YYAVCDPLRYKAKNIIIVICWIFISWSVPAVFAFGMIFLEINFGAEEIYKRVHCRG 181
 Db 121 YYAVCDPLRYKAKNIIIVICWIFISWSVPAVFAFGMIFLEINFGAEEIYKRVHCRG 180
 Oy 182 CSVFESKISGVLTFMTSFYIPGSIMLCVYRYIYLIAKEQARLISDANOKIQIGLEMKG 241
 Db 181 CSVFESKISGVLTFMTSFYIPGSIMLCVYRYIYLIAKEQARLISDANOKIQIGLEMKG 240
 Oy 242 SOSKERKAVKTLGIWGVFLICWCPFFICTVMDPFLHYIIPPTLNDVLIWFGYLNSTNP 301
 Db 241 SOSKERKAVKTLGIWGVFLICWCPFFICTVMDPFLHYIIPPTLNDVLIWFGYLNSTNP 300
 Oy 302 MYVAFFYPMFRKALXKMLFGKIFOKDSRCKLFLELS 339
 Db 301 MYVAFFYPMFRKALXKMLFGKIFOKDSRCKLFLELS 338
 XX
 RESULT 12
 ABB80694
 ID ABB80694 standard; protein; 339 AA.
 XX
 AC ABB80694;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human trace amine receptor.
 XX
 KW Biogenic receptor; trace amine receptor; synaptic transmission; human;
 KW hypertension; blood pressure; migraine; cardiac arrhythmia; seizure;
 KW coma; diabetes; schizophrenia; hypertension; asthma; drug addiction.
 XX
 OS Homo sapiens.
 XX

PN WO200222801-A2.
 XX
 PD 21-MAR-2002.
 XX
 PF 12-SEP-2001; 2001WO-US028455.
 XX
 PR 12-SEP-2000; 2000US-00659519.
 PR 09-JUL-2001; 2001US-0303967P.
 XX
 PA (UYOR-) UNIV OREGON HEALTH SCI.
 XX
 PI Bunzow JR, Grandy DK, Sonders M;
 XX
 DR WPI; 2002-371971/40.
 DR N-PSDB; ABL58316.
 XX
 PT A novel mammalian biogenic amine receptor, useful in assays to identify
 PT therapeutic agents for treating e.g., asthma and shock.
 XX
 PS Claim 4; Fig 1; 86pp; English.
 XX
 CC The invention relates to novel mammalian (human and rat) biogenic amine
 CC receptors. The mammalian biogenic receptors can be used to screen for
 CC compounds that bind to it, or to screen for compounds that inhibit a
 CC mammalian trace amine receptor. The compounds identified by the assay
 CC methods can be used to reduce sympathomimetic effects of enhanced trace
 CC amine dependent synaptic transmission in mammals so they can be used to
 CC to treat the peripheral effects of a drug that binds to or affects the
 CC binding to trace amine receptors e.g. hypertension, rapid heart rate,
 CC high blood pressure, migraine, cardiac arrhythmia, seizure, coma and
 CC diabetes or to treat pathological conditions associated with elevated
 CC levels of trace amines e.g. schizophrenia, depression, hypertension,
 CC shock, cardiac arrhythmias, asthma, migraine, psychosis, anaphylactic
 CC reactions and iatrogenic conditions. They are also useful for treating
 CC drug addiction. The present sequence represents a human trace amine
 CC receptor
 CC
 SQ Sequence 339 AA;
 XX
 Query Match 99.4%; Score 1791; DB 5; Length 339;
 Best Local Similarity 99.4%; Pred. No. 7,1e-195;
 Matches 337; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Oy 1 MPFCNININISCVKNNMSNDVRASLSLMVLIIITLVGNLIVVISISHFQKQHTPTNW 60
 Db 1 MPFCNININISCVKNNMSNDVRASLSLMVLIIITLVGNLIVVISISHFQKQHTPTNW 60
 Oy 61 IHSMAVDPLGCLVMPYSWVRSAEHCWYGEVFCIKHTSTIDIMLSASIFHLSFISID 120
 Db 61 IHSMAVDPLGCLVMPYSWVRSAEHCWYGEVFCIKHTSTIDIMLSASIFHLSFISID 120
 Oy 121 RYAVCDPLRYKAKNIIIVICWIFISWSVPAVFAFGMIFLEINFGAEEIYKRVHCRG 180
 Db 121 RYAVCDPLRYKAKNIIIVICWIFISWSVPAVFAFGMIFLEINFGAEEIYKRVHCRG 180
 Oy 181 GCSVFESKISGVLTFMTSFYIPGSIMLCVYRYIYLIAKEQARLISDANOKIQIGLEMKG 240
 Db 181 GCSVFESKISGVLTFMTSFYIPGSIMLCVYRYIYLIAKEQARLISDANOKIQIGLEMKG 240
 Oy 241 SOSKERKAVKTLGIWGVFLICWCPFFICTVMDPFLHYIIPPTLNDVLIWFGYLNSTNP 300
 Db 241 SOSKERKAVKTLGIWGVFLICWCPFFICTVMDPFLHYIIPPTLNDVLIWFGYLNSTNP 300
 Oy 301 PMVVAFFYPMFRKALXKMLFGKIFOKDSRCKLFLELS 339
 Db 301 PMVVAFFYPMFRKALXKMLFGKIFOKDSRCKLFLELS 339
 XX
 RESULT 13
 ADC12792
 ID ADC12792 standard; protein; 311 AA.
 XX
 AC ADC12792;
 XX

XX 18-DEC-2003 (first entry)
XX Human GPCR protein, SEQ ID NO 124.
XX
XX G protein-coupled receptor; GPCR; antibacterial; fungicide; protozoacide;
XX virocidic; antirheumatic; antiarthritic; tranquilliser; antidiabetic;
XX osteopathic; neurotropic; neuroprotective; anorectic; cardiant;
XX neuroleptic; cytostatic; antiparkinsonian; hypotensive; hypertensive;
XX antitumor; antiallergic; anticonvulsant; analgesic; infection; COPD;
XX rheumatoid arthritis; chronic obstructive pulmonary diseases; COPD;
XX asthma; non-insulin dependent diabetes; obesity; osteoporosis;
XX Alzheimer's disease; age-related macular degeneration;
XX myocardial infarction; schizophrenia; osteoarthritis; cancer;
XX Parkinson's disease; congestive heart failure; hypertension; hypertension;
XX ulcer; allergy; benign prostatic hyperplasia; seizure disorder; anxiety;
XX obsessive compulsive disorder; Cushing's syndrome; hypopituitarism; pain;
XX human.
XX
XX Homo sapiens.
XX
XX WO200300893-A2.
XX
XX 03-JAN-2003.
XX
XX 24-JUN-2002; 2002WO-IB002357.
XX
XX 26-JUN-2001; 2001US-0301095P.
XX 06-NOV-2001; 2001US-0333185P.
XX
XX (DECO-) DECODE GENETICS ENH.
XX
XX Martinez RMA, Sigurdsson GR;
XX
XX WPI, 2003-210155/20.
XX N-PSDB; ADC12791.
XX
XX New G protein-coupled receptor (GPCR) genes and polypeptides, useful for
XX diagnosing diseases associated with a GPCR, or in gene therapy for
XX treating e.g. obesity, osteoporosis, Alzheimer's, cancers or congestive
XX heart failure.
XX
XX Claim 10; SEQ ID NO 124; 253pp; English.
XX
XX The invention relates to a novel isolated nucleic acid of a G protein-
XX coupled receptor (GPCR) gene comprising any of 62 sequences of 912-2454
XX bp, or its complements; a GPCR polypeptide comprising any of 62 sequences
XX of 291-818 amino acids; or a nucleic acid that hybridises, under high
XX stringency conditions, with any of the 62 GPCR sequences or any of their
XX complements. The GPCR agents of the invention have the following
XX activities: antibacterial, fungicide, protozoacide, virucide,
XX antirheumatic, tranquiliser, antiarthritic, antidiabetic, osteopathic,
XX neurotropic, neuroprotective, anorectic, cardiant, neuroleptic, cytostatic,
XX antiparkinsonian, hypotensive, hypertensive, antitumor, antiallergic,
XX anticonvulsant, and analgesic. The GPCR therapeutic agent, particularly a
XX GPCR gene agonist or antagonist, is useful for treating a disease or
XX condition associated with a GPCR in an individual. The nucleic acid cited
XX above, which is 100 or fewer nucleotides in length, is useful for
XX assaying a sample for the presence of the GPCR gene nucleic acid or a
XX GPCR gene nucleic acid with at least one nucleotide difference from a
XX first nucleic acid, or for diagnosing a susceptibility to a disease or
XX conditions associated with a GPCR. These diseases include infections
XX (e.g. bacterial, fungal, protozoan or viral), rheumatoid arthritis,
XX chronic obstructive pulmonary diseases (COPD), asthma, non-insulin
XX dependent diabetes, obesity, osteoporosis, Alzheimer's disease, age-
XX related macular degeneration, myocardial infarction, schizophrenia,
XX osteoarthritis, cancers, Parkinson's diseases, congestive heart failure,
XX hypertension, hyperextension, ulcers, allergies, benign prostatic
XX hyperplasia, seizure disorder, anxiety, obsessive compulsive disorder,
XX Cushing's syndrome, hypopituitarism, or pain. This sequence represents
XX one of the 62 GPCR proteins of the invention.
XX
XX Sequence 311 AA;
XX
XX

Query Match 85.2%; Score 1539; DB 7; Length 311;
Beet Local Similarity 100.0%; Pred. No. 3.3e-166;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 29 LMLVLLITLVGNLIVISVSHFKQLHTPTNWLISHMATVDFLLGLVMPYSMAVSAHC 88
DB 23 LMLVLLITLVGNLIVISVSHFKQLHTPTNWLISHMATVDFLLGLVMPYSMAVSAHC 82
QY 89 MYFGVFEKIKTSVDIMLSSASIFLSTISIDRYAACDDPLRYKKNLIVYCNWIFLSW 148
DB 83 MYFGVFEKIKTSVDIMLSSASIFLSTISIDRYAACDDPLRYKKNLIVYCNWIFLSW 142
QY 149 SVPAVFAFGMTFLELNFKAEEIYKHHCHGCGCVFESKISGVLTFTMTSPYIPSGIMLC 208
DB 143 SVPAVFAFGMTFLELNFKAEEIYKHHCHGCGCVFESKISGVLTFTMTSPYIPSGIMLC 202
QY 209 VYRYIYLAKQARLISDANQKLGLEMKNGISQSKERKAVKTLGIWGVFLICWCPFF 268
DB 203 VYRYIYLAKQARLISDANQKLGLEMKNGISQSKERKAVKTLGIWGVFLICWCPFF 262
QY 269 ICTVWDPEPLHYITPPTLNDVLIWPGYLNSTENPMVYAFYEMFRKALM 317
DB 263 ICTVWDPEPLHYITPPTLNDVLIWPGYLNSTENPMVYAFYEMFRKALM 311
RESULT 14
AAG80972 ID AAG80972 standard; protein, 296 AA.
XX
XX AAG80972;
XX
XX 28-AUG-2001 (first entry)
XX
XX Human nGPCR56 #2.
XX
XX G protein-coupled receptor; nGPCR; seven transmembrane receptor;
XX signal transduction; schizophrenia; thyroid disorder; renal failure;
XX rheumatoid arthritis; CNS disorder; infection; metabolic disease;
XX cardiovascular disease; proliferative disorder; hormonal disorder;
XX neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
XX attention deficit-hyperactivity disorder/attention deficit disorder;
XX Parkinson's disease; migraine; senile dementia; inflammatory disease;
XX rheumatoid arthritis; autoimmune disorder; respiratory ailment;
XX neuroprotective.
XX
XX Homo sapiens.
XX
XX WO200136473-A2.
XX
XX 25-MAY-2001.
XX
XX 16-NOV-2000; 2000WO-US031581.
XX
XX 16-NOV-1999; 99US-0165838P.
XX 17-NOV-1999; 99US-0166071P.
XX 19-NOV-1999; 99US-016678P.
XX 28-DEC-1999; 99US-0173396P.
XX 22-FEB-2000; 2000US-0184129P.
XX 28-FEB-2000; 2000US-0185421P.
XX 28-FEB-2000; 2000US-0185549P.
XX 02-MAR-2000; 2000US-0186530P.
XX 03-MAR-2000; 2000US-0186811P.
XX 09-MAR-2000; 2000US-0188119P.
XX 17-MAR-2000; 2000US-0190310P.
XX 21-MAR-2000; 2000US-0190800P.
XX 20-APR-2000; 2000US-0198568P.
XX 02-MAY-2000; 2000US-0201190P.
XX 08-MAY-2000; 2000US-0203111P.
XX 25-MAY-2000; 2000US-0207094P.
XX
XX (PRAA) PHARMACIA & UPJOHN CO.
XX
XX

PI Vogel I G, Wood LS, Parodi LA, Hiebach RR, Lind P, Slighcom J,
PI Schellin KA, Kayles PS, Bannigan CW, Ruff V, Sejltz T, Huff RM,
XX
DR WPI: 2001-389826/41.
XX N-PSDB; AAH51012.
PT New G protein-coupled receptor (ngPCR-x) and its encoding polynucleotide
PT useful for diagnosing and treating e.g. schizophrenia.
XX
PS Claim 37; Page 91-92; 261pp; English.
XX
XX The present invention relates to novel G protein-coupled receptors
CC (ngPCR; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27, 28,
CC 31-38, 40, 41, 53-60) and their coding sequences. The present sequence is
CC one such G protein-coupled receptor. GPCRs are also known as seven
CC transmembrane receptors and function in signal transduction. The ngPCR
CC coding sequences are useful for screening a human to diagnose a disorder
CC affecting the brain or a genetic predisposition, specifically
CC schizophrenia. ngPCR are useful for identifying compounds useful for
CC treating schizophrenia. Detection of ngPCR in a sample is useful as a
CC diagnostic tool for diseases or disorders e.g. thyroid disorders, renal
CC failure, rheumatoid arthritis, CNS disorders, infections such as HIV-1,
CC metabolic and cardiovascular diseases, proliferative disorders and
CC hormonal disorders. Modulators of ngPCR activity have the utility for
CC treating neurological disorders, including schizophrenia, ADHD/ADD
CC (attention deficit-hyperactivity disorder/attention deficit disorder),
CC and neuronal disorders such as Alzheimer's disease, Parkinson's disease,
CC migraine and senile dementia. Additional disorders include inflammatory
CC conditions (e.g. Crohn's disease), rheumatoid arthritis, autoimmune
CC disorders, cancers, respiratory ailments such as asthma, and inflammatory
CC diseases e.g. inflammatory bowel disease
XX
SQ Sequence 296 AA;

Query Match 84.4%; Score 1525; DB 4; Length 296;
Best Local Similarity 100.0%; Pred. No. 1.2e-164;
Matches 287; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNPFCNNIINISCVKNNNSNDVRASLYSLMVLIIITLVGNLIVIVISHPKQLHTPTNW 60
DB 1 MNPFCNNIINISCVKNNNSNDVRASLYSLMVLIIITLVGNLIVIVISHPKQLHTPTNW 60
QY 61 LIHSNATVDFLLGLCLVMPYSWVSAEHWCYFGEVFCIKHTSTDIMLSASIFHLSFISID 120
DB 61 LIHSNATVDFLLGLCLVMPYSWVSAEHWCYFGEVFCIKHTSTDIMLSASIFHLSFISID 120
QY 121 RYVAUCDPLRYAKNNIIVICWIFISWSVPAVFAFGMI FLELNFGAEIYYKVVHCRG 180
DB 121 RYVAUCDPLRYAKNNIIVICWIFISWSVPAVFAFGMI FLELNFGAEIYYKVVHCRG 180
QY 181 GGSVPSKISGVLTPMTSFYIPGSIIMLCVYRIRIYLAKEQARLISDANOKLQIGLEMKNG 240
DB 181 GGSVPSKISGVLTPMTSFYIPGSIIMLCVYRIRIYLAKEQARLISDANOKLQIGLEMKNG 240
QY 241 ISQSKERKAVKTLGIYMGVFLICWCPFFICTYMDPFLHYIIPPTLND 287
DB 241 ISQSKERKAVKTLGIYMGVFLICWCPFFICTYMDPFLHYIIPPTLND 287
RESULT 15
ABG93790
ID ABG93790 standard; protein; 296 AA.
XX
AC ABG93790;
XX
XX 26-NOV-2002 (first entry)
XX
DE Human G protein-coupled receptor protein, ngPCR-56, #1.
XX
XX Human, receptor; G protein-coupled receptor; GPCR; ngPCR; beGPCR;
KW ng protein coupled receptor; communication; serpentine structure;
KW seven transmembrane receptor; 7TM; mental disorder; diagnosis;
KW genetic predisposition; brain; immune response; gene therapy;

KW anxiety disorder; depression; bipolar disorder; schizophrenia;
KW Huntington's disease; dyskinesia; manic depression; stroke;
KW Parkinson's disease; Alzheimer's disease; diabetes; inflammation; wound;
KW tranquiliser.
XX
OS Homo sapiens.
XX
PN WO200264789-A1.
XX
XX 22-AUG-2002.
XX
XX 14-FEB-2001; 2001WO-US004641.
XX
XX 14-FEB-2001; 2001WO-US004641.
XX
XX (PHARMA) PHARMACIA & UPROHN CO.
XX
PI Lind P, Parodi LA, Vogel I G, Wood LS;
XX
DR WPI: 2002-674879/72.
XX N-PSDB; AB870245.
XX
XX New nucleic acids and polypeptides of the ng protein-coupled receptor,
PT useful for treating or diagnosing a mental disorder or a disorder
PT affecting the brain, e.g. anxiety disorders, schizophrenia, stroke or
PT Parkinson's disease.
XX
PS Example 1; Page 86-87; 244pp; English.
XX
XX The invention discloses an isolated human polypeptide, and encoding
CC nucleic acid, for a G protein-coupled receptor (GPCR), particularly the
CC ng protein coupled receptor-14 (ngPCR-14). GPCRs are vital in the
CC communication between cells and their environment and are characterised
CC by a serpentine structure that passes through the cell membrane seven
CC times, hence the reason such receptors are sometimes called seven
CC transmembrane receptors (7TM). The polynucleotides and polypeptides are
CC useful for identifying an ngPCR allelic variant that correlates with a
CC mental disorder, for isolating an antibody that binds to an epitope of
CC the polypeptide, for identifying a compound that binds the polypeptide or
CC polynucleotide and/or modulates its biological activity, for screening a
CC human subject to diagnose a disorder, or a genetic predisposition to a
CC disorder, affecting the brain or a genetic disposition to the disorder,
CC for identifying compounds useful for the treatment of a mental disorder
CC and for identifying a compound useful as a modulator of binding between
CC ngPCR-14 and a binding partner of ngPCR-14. The polypeptide is also
CC useful for inducing an immune response in a mammal. The nucleic acid or
CC polypeptide is particularly useful, using gene therapy, for treating e.g.
CC anxiety disorders, depression, bipolar disorder, schizophrenia,
CC Huntington's disease, dyskinesias, manic depression, stroke, Parkinson's
CC disease or Alzheimer's disease. The nucleic acid and polypeptide may also
CC be used for treating diabetes, inflammation or wounds. The sequences
CC presented in ABG93747-ABG93793, ABG93795 and ABG93796 are the ngPCR (also
CC referred to as beGPCR) proteins
XX
SQ Sequence 296 AA;

Query Match 84.4%; Score 1525; DB 5; Length 296;
Best Local Similarity 100.0%; Pred. No. 1.2e-164;
Matches 287; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNPFCNNIINISCVKNNNSNDVRASLYSLMVLIIITLVGNLIVIVISHPKQLHTPTNW 60
DB 1 MNPFCNNIINISCVKNNNSNDVRASLYSLMVLIIITLVGNLIVIVISHPKQLHTPTNW 60
QY 61 LIHSNATVDFLLGLCLVMPYSWVSAEHWCYFGEVFCIKHTSTDIMLSASIFHLSFISID 120
DB 61 LIHSNATVDFLLGLCLVMPYSWVSAEHWCYFGEVFCIKHTSTDIMLSASIFHLSFISID 120
QY 121 RYVAUCDPLRYAKNNIIVICWIFISWSVPAVFAFGMI FLELNFGAEIYYKVVHCRG 180
DB 121 RYVAUCDPLRYAKNNIIVICWIFISWSVPAVFAFGMI FLELNFGAEIYYKVVHCRG 180
QY 181 GGSVPSKISGVLTPMTSFYIPGSIIMLCVYRIRIYLAKEQARLISDANOKLQIGLEMKNG 240
DB 181 GGSVPSKISGVLTPMTSFYIPGSIIMLCVYRIRIYLAKEQARLISDANOKLQIGLEMKNG 240

Db	181	GCSEFFSKISGVLFMTISFYIPGSIMLCVYRITLAKAQARLISDANOQLQIGLEMKNG	240
Qy	241	ISQSKERKAVKTLGIWGVFLICWCPFFICTVMDPFLHYIIPPTLND	287
Db	241	ISQSKERKAVKTLGIWGVFLICWCPFFICTVMDPFLHYIIPPTLND	287

Search completed: February 12, 2005, 03:43:20
Job time : 78 secs


```

; Sequence 4, Application US/09659519
; Patent No. 6783973
; GENERAL INFORMATION:
; APPLICANT: Grandy, David K
; APPLICANT: Bunzow, James R
; TITLE OF INVENTION: Mammalian Catecholamine Receptor Genes and Uses
; FILE REFERENCE: Catecholamine receptor genes
; CURRENT APPLICATION NUMBER: US/09/659,519
; CURRENT FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-09-659-519-4

```

```

Query Match      77.1%; Score 1393; DB 4; Length 332;
Best Local Similarity 77.5%; Pred. No. 1,2e-114;
Matches 259; Conservative 27; Mismatches 46; Indels 2; Gaps 1;

```

```

QY 2 MPECHNIINISCVKNNMSNDVRASLSLWVLIITLTGNTLIVYISIFSKQIHTPTNWL 61
DB 1 MHLCHSANISHNRMSRPDRASLSLSLILTLVGNLIVYISIFSKQIHTPTNWL 60
QY 62 IHSMTVDPLGLVMPYSWVRSAEHCWYFGEVFCIKHTSTDMLSASIFHLSFISDR 121
DB 61 LHSMAVVDPLGLVMPYSWVRSAEHCWYFGEVFCIKHTSTDMLSASIFHLSFISDR 120
QY 122 YVAVCDPLLYKAKMNLIVICWIFISWSPVPAFAMITLLENFKAETIYKHYHCRG 181
DB 121 YVAVCDPLLYKAKMNLIVICWIFISWSPVPAFAMITLLENFKAETIYKHYHCRG 180
QY 182 CSVFESKISGVLTFTMTSPYIPGSIIMLCVYRYIYLAKEOARLISDANOKLOIGLEMKGI 241
DB 181 CFLFESKVGVLAFMTSPYIPGSIIMLCVYRYIYLAKEOARLISDANOKLOIGLEMKGI 238
QY 242 SCSKRAVKITIGIYWGVLICWCPFFICTVMDPLHYIIPPTLNDVLIWFGYLNSTNP 301
DB 239 PSKSKTKAKITIGIYWGVLICWCPFFICTVMDPLHYIIPPTLNDVLIWFGYLNSTNP 298
QY 302 MYVAFYFWFRKALKMFLGKIFQKDSRCKLFL 335
DB 299 MYVAFYFWFRKALKMFLGKIFQKDSRCKLFL 332

```

RESULT 3

```

US-08-788-539A-2
; Sequence 2, Application US/08788539A
; Patent No. 5871967

```

GENERAL INFORMATION:

```

; APPLICANT: SmithKline Beecham Corporation
; TITLE OF INVENTION: CLONING OF A NOVEL G-PROTEIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/788,539A
; FILING DATE: 24-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:

```

```

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: ATG50047
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-4026
;
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 343 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-788-539A-2

```

```

Query Match      45.0%; Score 813.5; DB 2; Length 343;
Best Local Similarity 46.0%; Pred. No. 1e-63;
Matches 159; Conservative 57; Mismatches 99; Indels 31; Gaps 5;

```

```

QY 5 CHNIN-----ISCVKNN-----MSNDVRASLSLWVLIITLTGNTLIVYISISH 50
DB 13 CPKFNKILSHQPLFSCPDGNVFGYDKSHDY-----PLFGNLIVYISISH 58
QY 51 FKOIHTPTNWL IHSMTVDPLGLVMPYSWVRSAEHCWYFGEVFCIKHTSTDMLSAS 110
DB 59 FKOLHSPNPLILSMATDPLGLVMPYSWVRSAEHCWYFGEVFCIKHTSTDMLSAS 118
QY 111 IFHLSFISDRYVAVCDPLLYKAKMNLIVICWIFISWSPVPAFAMITLLENFKAETI 170
DB 119 IFHLSFISDRYVAVCDPLLYKAKMNLIVICWIFISWSPVPAFAMITLLENFKAETI 178
QY 171 IYVHVHCRGGSVFSKISGVLTFTMTSPYIPGSIIMLCVYRYIYLAKEOARLISDANOK 230
DB 179 -YKILVACFNFCALTFFKFWGTILFTTCFTPGSIIMVIGIKFIYVSKQARVISHPEN 237
QY 231 LOIGLEMKNGISCSKRAVKITIGIYWGVLICWCPFFICTVMDPLHYIIPPTLNDVLI 290
DB 238 TKGA--YKHLSSKQDKAKKITIGIYWGVLICWCPFFICTVMDPLHYIIPPTLNDVLI 295
QY 291 WFGYLNSTNP MYVAFYFWFRKALKMFLGKIFQKDSRCKLFL 336
DB 296 WLRFYNSTGNPLIHGFNFWFKAKFYIVSGKIFSSHSEITANLPFE 341

```

RESULT 4

```

US-09-286-805-4
; Sequence 4, Application US/09286805
; Patent No. 6117990

```

GENERAL INFORMATION:

```

; APPLICANT: Bontini, James A.
; APPLICANT: Borowsky, Beth E.
; TITLE OF INVENTION: DNA Encoding Orphan SNORF1 Receptor
; FILE REFERENCE: 58987
; CURRENT APPLICATION NUMBER: US/09/286,805
; CURRENT FILING DATE: 1999-04-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 4
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-09-286-805-4

```

```

Query Match      43.3%; Score 782; DB 3; Length 338;
Best Local Similarity 45.8%; Pred. No. 5.9e-61;
Matches 154; Conservative 63; Mismatches 115; Indels 4; Gaps 4;

```

```

QY 2 MPECHNIINISCVKNNMSNDVRASLSLWVLIITLTGNTLIVYISIFSKQIHTPTNWL 61
DB 1 MELCTENVNGSCIKSSYSFMPRAIIYAVVIGALLAVGNNLIVTAIHFKQIHTPTNWL 60

```

[illegible]

```

RESULT 5
US-09-286-805-2
: Sequence 2, Application US/09286805
: Patent No. 617990
: GENERAL INFORMATION:
: APPLICANT: Bonini, James A.
: APPLICANT: Borovsky, Beth E.
: TITLE OF INVENTION: DNA Encoding Orphan SNORF1 Receptor
: FILE REFERENCE: 58967
: CURRENT APPLICATION NUMBER: US/09/286, 805
: CURRENT FILING DATE: 1999-04-06
: NUMBER OF SEQ. ID NOS.: 4
: SOFTWARE: PatentIn Ver. 2.0 - beta
: SEQ. ID NO. 2
: LENGTH: 348
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-286-805-2

```

	Query Match	41.1%	Score 743	DB 3	Length 348	
	Best Local Similarity	42.5%	Pred 1.6e-57			
	Matches 141	Conservative	63	Mismatches 126	Indels 2	Gaps 2
QY	5	CHNIINISCVKNNMSNDVRA	SLYSIMLVIIILTTLVGNLIVISISHFKQLHTPTNMLIHS	64		
Db	14	CYKNNVESCITKTPYSPGPR	SLIYAVLGGGAVLAARGNLLMALHFKQLHTPTNMLIAS	73		
QY	65	MATVDFELGCLVMPYSWM	RAEHCMYFGEVFCCKHTSTDIMLSASIFHLSTSIDRYA	124		
Db	74	LACADFVLGVATVMPST	RSVSESCMYFGDPSYCKFHFTCEPDSICFPASLFLHCCISIDRYA	133		
QY	125	VCDPLRYAKAKNNIIIVC	MIIFSMGVPVAFACGMFLDLNFKGAREIYKHHGCGGCV	184		
Db	134	VTDLPTTPTKKTTSVSG	ICIVLSWFFSVYTSISITYTGANESGIEBLVVA-LTCVGGCOA	192		
QY	185	FFSKISGVLTFMTSEYIP	SGIMLCVYRIYLIYLAQKQARLISDANOKLQIGLEMKRGISOS	244		
Db	193	PLNQNMWVLICFL-FF	INVMVFLYSKIFLVAKQARKIESTASQAQSSSYKERVAK	251		
QY	245	KERAAVKTGLIVMGVFL	ICWCPFICTVMDPEFLHYIIPPTLNDVLIWFGYLNSTNPVY	304		
Db	252	RERRAKTLGIAMAFAFL	VMFLPYLDAVIDAMNFTPTPYVEIILVMCVYNSAMNPITY	311		
QY	305	AFPYWPKALKMLFGKI	FOQDSRCULFLE	336		
Db	312	AFPYQWFGKAIKLIVS	GKVLRTDSSSTTNLFSE	343		

RESULT 6
US-08-467-559B-2
; Sequence 2, Application US/08467559B

Patent No.5928690
GENERAL INFORMATION:
APPLICANT: LI, YI
TITLE OF INVENTION: HUMAN AMINE RECEPTOR
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN AND FOX, P.L.L.C.
STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: UNITED STATES OF AMERICA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,559B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K
REGISTRATION NUMBER: 36, 688
REFERENCE/DOCKET NUMBER: 1488.0840000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 337 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

Query Match	Similarity	37.5%	Score 677.5	DB 2	length 337
Best Local	Similarity 39.0%	Pred. No. 9e-52			
Matches 130	Conservative 67	Mismatches 125	Indels 11	Gaps 4	
Qy	4	FCHNIINISCVKNNMSNDVRASLYSLMWLLITLTVLGNLIIVISISHEFKOLHTPTNNLHI	63		
Db	16	FCYQ-VANGSCPRVHTHTIGLVLTITLCAAGMLIIVLGNVFAFANYSYKALHTPTNFPILL	74		
Qy	64	SMATVDELGLCLVMPYMSVRSAEHCWYFGEVCKIHTSTDMLSASIFHLSFISIDRY	123		
Db	75	SLALADNFIGLTVLPLSTIRSVESCMFEGDELCRLHTYLDLFCLTSSIFHLCEFISIDBHC	134		
Qy	124	AVCPPLAYKAKMNLIVICWLFISMSVAVFAFGMIPELNFKGAEEIYKRVHVRGGGS	183		
Db	135	AIICPPLIYPSKFYTRVALRYTLAAGVPAATYSLEFLYDVVETRISQ-WLEEMPVCSQ	193		
Qy	184	VFPFKISGLVLFMTSFYIPGSIMLCVYRIYLIAKEQRLISDANKQIQIGLEMKGISQ	243		
Db	194	ILLNKFPGWLNLF-PLFVPCILIMSLYKIRIVAVTRQOQITLTSKSL-----AGAA	244		
Qy	244	SKERRAAVKTIGIVNGVFLICWCPFFICTVMDPELHYIIPPTLNDVLINFGYINSTFNNPV	303		
Db	245	KHEERAKTIGIVGIYDLCLWLPETIDMVSLHFIFNPVDFIIFWAFYENSACNPIT	304		
Qy	304	YAFYPMFRKLKRMFLPGKIFQKXSSRCKLPLE	336		
Db	305	YFYSYQWFRKALKLTLSQKVSPPQRTIVDAOE	337		

```

RESULT 7
US-09-555-1133-4
: Sequence 4, Application US/09555313B
: Patent No. 6506580
: GENERAL INFORMATION:
: APPLICANT: FISCHWEISTER, Rudolph et al.
: TITLE OF INVENTION: Splicing variants of the human serotoninergic receptor
: TITLE OF INVENTION: 5-HT4 and uses thereof, in particular for screening

```

```
FILE REFERENCE: P06762US00/BAS
CURRENT APPLICATION NUMBER: US/09/555,313B
CURRENT FILING DATE: 2002-08-13
PRIOR APPLICATION NUMBER: FR 97/15037
PRIOR FILING DATE: 1997-11-28
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 4
LENGTH: 360
TYPE: PRT
ORGANISM: Homo sapiens
US-09-555-313B-4
```

```
Query Match      29.6%; Score 534; DB 4; Length 360;
Best Local Similarity 37.4%; Pred. No. 3.9e-39;
Matches 116; Conservative 53; Mismatches 121; Indels 20; Gaps 5;
```

```

26 LYSIMVLIIITLVGNLIIVISISHPKQLH-TPTNMLIHSMTAVDFLGLCLVMPYSWRS 84
22 LITFLSTVILMALIGNLVMAVAVCMWRQLRKIKTNFYISLAFADLVSLVWPFGAIEL 81
85 AEHCWYFGEVFCIKHTSTDIIMLSASIFHLFSISIDRYAV-CDPLRYKAKNNILVICWM 143
82 VQDIWIYGEVFCIVRSISLVLTITASI FHLCCISIDRYAICCPLYVRNKKMTPLRIALM 141
144 IFTISNVPVAFVAFGMIFELNFKGAEIYYKHVHCRG-----CSVFESKISGVLTFMTSF 199
142 LGGCWVIFPTFISFLPMOGMNNIGIIDLEKRFNONSNTYCVPMVWKPAITCSVAAF 201
200 YIPGSIIMLCVYRYIYLIAKEQARLISDANOKLOIGLEMKNGISQS-----KERKA 249
202 YIPFLMLVATYRIYVTAKEHAHQI---QMLQAGASSESRPSADQSTHRMRTETKA 257
250 VKTLGIWGVFLICWCPFFICTVMDPFLHYIIPETLNDVLIWFGYLNSTFNPVYAFEPY 309
258 AKTLCTIIMGFCICWAPFFVINIVDFIDYVPGQWTAFLMLGYSGLNPLFYAFPLNK 317
310 WFRKRLKML 319
318 SFRRAFLIIL 327
```

```
RESULT 8
US-09-555-313B-24
Sequence 24, Application US/09555313B
Patent No. 6506580
GENERAL INFORMATION:
APPLICANT: FICSMIEISTER, Rudolph et al.
TITLE OF INVENTION: Splicing variants of the human neurotrophic receptor
FILE REFERENCE: P06762US00/BAS
CURRENT APPLICATION NUMBER: US/09/555,313B
CURRENT FILING DATE: 2002-08-13
PRIOR APPLICATION NUMBER: FR 97/15037
PRIOR FILING DATE: 1997-11-28
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 24
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-555-313B-24
```

```
Query Match      29.6%; Score 534; DB 4; Length 380;
Best Local Similarity 37.4%; Pred. No. 4.2e-39;
Matches 116; Conservative 53; Mismatches 121; Indels 20; Gaps 5;
```

```

26 LYSIMVLIIITLVGNLIIVISISHPKQLH-TPTNMLIHSMTAVDFLGLCLVMPYSWRS 84
22 LITFLSTVILMALIGNLVMAVAVCMWRQLRKIKTNFYISLAFADLVSLVWPFGAIEL 81
85 AEHCWYFGEVFCIKHTSTDIIMLSASIFHLFSISIDRYAV-CDPLRYKAKNNILVICWM 143
```

```

82 VQDIWIYGEVFCIVRTSISLVLTITASI FHLCCISIDRYAICCPLYVRNKKMTPLRIALM 141
144 IFTISNVPVAFVAFGMIFELNFKGAEIYYKHVHCRG-----CSVFESKISGVLTFMTSF 199
142 LGGCWVIFPTFISFLPMOGMNNIGIIDLEKRFNONSNTYCVPMVWKPAITCSVAAF 201
200 YIPGSIIMLCVYRYIYLIAKEQARLISDANOKLOIGLEMKNGISQS-----KERKA 249
202 YIPFLMLVATYRIYVTAKEHAHQI---QMLQAGASSESRPSADQSTHRMRTETKA 257
250 VKTLGIWGVFLICWCPFFICTVMDPFLHYIIPETLNDVLIWFGYLNSTFNPVYAFEPY 309
258 AKTLCTIIMGFCICWAPFFVINIVDFIDYVPGQWTAFLMLGYSGLNPLFYAFPLNK 317
310 WFRKRLKML 319
318 SFRRAFLIIL 327
```

```
RESULT 9
US-08-446-822-8
Sequence 8, Application US/08446822
Patent No. 5766879
GENERAL INFORMATION:
APPLICANT: SYNAPTIC PHARMACEUTICAL CORPORATION
TITLE OF INVENTION: DNA ENCODING 5-HT4 SEROTONIN RECEPTORS
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: COOPER & DUNHAM
STREET: 30 ROCKEFELLER PLAZA
CITY: NEW YORK
STATE: NEW YORK
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,822
FILING DATE: June 1, 1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: White, P., John
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 42667-A-PCT-US/JPW/MAT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 388 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-446-822-8
```

```
Query Match      29.6%; Score 534; DB 1; Length 388;
Best Local Similarity 37.4%; Pred. No. 4.3e-39;
Matches 116; Conservative 53; Mismatches 121; Indels 20; Gaps 5;
```

```

26 LYSIMVLIIITLVGNLIIVISISHPKQLH-TPTNMLIHSMTAVDFLGLCLVMPYSWRS 84
22 LITFLSTVILMALIGNLVMAVAVCMWRQLRKIKTNFYISLAFADLVSLVWPFGAIEL 81
85 AEHCWYFGEVFCIKHTSTDIIMLSASIFHLFSISIDRYAV-CDPLRYKAKNNILVICWM 143
82 VQDIWIYGEVFCIVRTSISLVLTITASI FHLCCISIDRYAICCPLYVRNKKMTPLRIALM 141
144 IFTISNVPVAFVAFGMIFELNFKGAEIYYKHVHCRG-----CSVFESKISGVLTFMTSF 199
142 LGGCWVIFPTFISFLPMOGMNNIGIIDLEKRFNONSNTYCVPMVWKPAITCSVAAF 201
```



```

; FILE REFERENCE: P06762US00/BAS
; CURRENT APPLICATION NUMBER: US/09/555,313B
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: FR 97/15037
; PRIOR FILING DATE: 1997-11-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-555-313B--2

```

Query Match 29.6%; Score 534; DB 4; Length 760;
Best Local Similarity 37.4%; Pred. No. 8.9e-39;
Matches 116; Conservative 53; Mismatches 121; Indels 20; Gaps 5;

QY 26 LYSIMALLITTTTGLNLIIVISISHKQCH-TPTWMLISMATVDFFLGGTLVMPSSMVR84
Db 22 LLPLFLSTVLIMALIGNLVMAVCWDRLRKIKNTNYFVLSLAFLADLVSVSLVMPFGAILEI81

Dy
85 AAEHCWTFGFEVFCIKHTSDIMLSASIFILHSFISIDRYAAV-CGPLRYAKNNLIVICWM 143

Db
82 VQDIMIYGEVEFLCVRTSLDVLITLASIPHLCCSIDRYYAICCPLVYNKMTPELRIAM 141

Db
144 I F S M S V A A V F A R G M I L D E L N F K A E E I Y K H V H C R G G --- C S V F S K I S G V L T F M E S F 199
142 I G G C M V A I P T F I S F L P I M O G M N N I G I I D L E K R K N Q N S N S T Y C V F M A N K E Y A I T C S V A F 201

Db
202 YIPLLMLLAYRIRIVTAKHAHQI---QMLGRAGASSESRPGSADQSHTRHRTETKA 257

258 ATTCCTGCGCCTCWAPEFVNINIVDPFDITYTGQWVTAFLWIGYINSGINPELYAFLNK 317

Db ||:| ::|
318 SRRRAFLIL 327

RESULT 13
US-09-826-509-443
: Sequence 443, Application US/09826509

```

; GENERAL INFORMATION:
; APPLICANT: Lehmann-Brulnsma, Karin
; APPLICANT: Liaw, Chen W.
; APPLICANT:

```

TITLE OF INVENTION: Protein-Coupled Receptors
 TITLE OF INVENTION: Constitutively Activated Known G
 NO. 6806054-Endogenous,
 FILE REFERENCE: AREN-207
 CURRENT APPLICATION NUMBER: US/06/0936 E00

PRIOR FILING DATE: 2001-04-05
PRIORITY APPLICATION NUMBER: 60/195,747
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: 09/170,496

```

1  NUMBER OF SEQ ID NOS: 589
2  SOFTWARE: PatentIn Version 2.1
3  SEQ ID NO 443

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-826-509-443

```

Query Match	29.5%	Score 532,	DB 4,	Length 360,
Best Local Similarity	37.4%	Pred. No. 5.9e-39,		
Matches 116; Conservative	53;	Mismatches 121;	Indels 20;	Gaps 5;

26 LKSLAVLLITLVNGLIVIVSISHKQLH-TPNNLIHSNATVDLIGGLVMPYSMVR 84

Db 22 LITPSTVITAMIALGNLWVAVCMWDGRQKRIKINVEIYSLAFADLLVSVLWVPQAIEL 81

Qy 85 AEHCWYFEGVFECKIHSTDIMLSASIFHLSTSIDRYAV-CDPIRYAKONILVICYW 143

Db 82 VDDIMYGEVFLVRISLDVLLTTTASIFHLCCSIDRYAALICOPLVYNNKMTPIRIALM 141

Qy 144 IFISWSPVAFPMGMI FLEINFKGAEIYYKGVHCKG- - -CSVEFSKISGLVTMTSF 199

Db 142 LGGCWVPIPFISFLPIMGOMNIGIIDLEKRRKNONSSTYCVPMVNNPYAITSVAF 201

Qy 200 YIPGSIIMLCVYRIYLIKAQEARLISDANOKIOTIGLEMKNGISQ- - - - -KERQA 249

Db 202 YIPFLIMLAVYRIYVTAKEHNAQ- - - - -QMQRAGASSESPQSDOSHTRHMRJETYA 257

Qy 250 VETLIGIWGVFLICWCPFEICTVMDPFLYHIIPIPLINDLIRFGYLSSTFENMVVAFEP 309

Db 258 KTLICIIIGCFCLCWAPFVTNIYDPIDYTVPGVMTAFWLGITNSGIANFLVAFMK 317

Qy 310 WERKALKMML 319

Db 318 SFRRAFLIIL 327

RESULT 14
US-09-826-509-445

```

; Patent No. 6806054
;
; GENERAL INFORMATION:
; APPLICANT: Lehmann-Brunsma, Karin
; APPLICANT: Frau Christa

```

APPLICANT: LIN, I-MIN
TITLE OF INVENTION: No. 6806054-Endogenous, Constitutively Activated Known
TITLE OF INVENTION: Protein-coupled Receptors
FILE REFERENCE: APEN-207

```

CURRENT FILING DATE: 06/07/2009,309
CURRENT FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/195,747
PRIOR FILING DATE: 2000-04-07

```

```

; PRIOR FILING DATE: 1998-10-13
;
; NUMBER OF SEQ ID NOS: 589
;
; SOFTWARE: PatentIn Version 2.1

```

```

;      LENGTH: 378
;      TYPE: PRT
;      ORGANISM: Homo sapiens
;

```

Query Match	29.5%;	Score 532;	DB 4;	Length 378;
Best Local Similarity	37.4%;	Pred. No. 6;	2e-39;	

26 LKSLMVLILTLVGNLIVISISHFKOLH-PTNNWLHSMATVDFLLGLVMPYSMRS 84

```

85 AHHCWYGEVFCFKIHTSTDIMLSSASIFHLFSIDRYAV-CDPLRYAKONITLVICM 143
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
82 WDNIATVGEVETVETETDVIETETETETETETETETETETETETETETETETETETET
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

QY 144 IFISMSVPAVFAFGMIFLELNFGAEIEIYKXHVHCRG---CSVFESKISGVLTFMTSTF 199

142 IGGCWVLPETISELPITMGQNNITGITITREPREKNNNSGATVCTPMNIPDVTTCGGTATF 201

[illegible]

Qy 250 VKTLGIAMGVLLICWCBPFICTVMDPELHYIIPTTLNDVLILFMFGYLNSTFENMVYAEFP 309
258 KTLICIMGCFCOLMAPEFVNIVDEPIDYATPGOWTAPILMIGYTSGLNPTLYAFINLK 317
zb

QY 310 WPKALJOML 319

Db 22 LITPSTVITAMIALGNLWVAVCMWDGRQKRIKINVEIYSLAFADLLVSVLWVPQAIEL 81

Qy 85 AEHCWYFEGVFECKIHSTDIMLSASIFHLSTSIDRYAV-CDPIRYAKONILVICYW 143

Db 82 VDDIMYGEVFLVRISLDVLLTTTASIFHLCCSIDRYAALICOPLVYNNKMTPIRIALM 141

Qy 144 IFISWSPVAFPMGMI FLEINFKGAEIYYKGVHCKG- - -CSVEFSKISGLVTMTSF 199

Db 142 LGGCWVPIPFISFLPIMGOMNIGIIDLEKRRKNONSSTYCVPMVNNPYAITSVAF 201

Qy 200 YIPGSIIMLCVYRIYLIKAQEARLISDANOKIOTIGLEMKNGISQ- - - - -KERQA 249

Db 202 YIPFLIMLAVYRIYVTAKEHNAQ- - - - -QMQRAGASSESPQSDOSHTRHMRJETYA 257

Qy 250 VETLIGIWGVFLICWCPFEICTVMDPFLYHIIPIPLINDLIRFGYLSSTFENMVVAFEP 309

Db 258 KTLICIIIGCFCLCWAPFVTNIYDPIDYTVPGVMTAFWLGITNSGIANFLVAFMK 317

Qy 310 WERKALKMML 319

Db 318 SFRRAFLIIL 327

RESULT 14
US-09-826-509-445

```

; Patent No. 6806054
;
; GENERAL INFORMATION:
; APPLICANT: Lehmann-Brunsma, Karin
; APPLICANT: Frau Christa

```

APPLICANT: LIN, I-MIN
TITLE OF INVENTION: No. 6806054-Endogenous, Constitutively Activated Known
TITLE OF INVENTION: Protein-coupled Receptors
FILE REFERENCE: APEN-207

```

CURRENT FILING DATE: 06/07/2009, 309
CURRENT FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/195,747
PRIOR FILING DATE: 2000-04-07

```

```

; PRIOR FILING DATE: 1998-10-13
;
; NUMBER OF SEQ ID NOS: 589
;
; SOFTWARE: PatentIn Version 2.1

```

```

;      LENGTH: 378
;      TYPE: PRT
;      ORGANISM: Homo sapiens
;

```

Query Match	29.5%;	Score 532;	DB 4;	Length 378;
Best Local Similarity	37.4%;	Pred. No. 6;	2e-39;	

26 LKSLMVLILTLVGNLIVISISHFKOLH-PTNNWLHSMATVDFLLGLVMPYSMRS 84

```

85 AHHCWYGEVFCFKIHTSTDIMLSSASIFHLFSISIDRYAV-CDPLRYAKONITLVICVM 143
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
82 WDNIATVGEVETVETETDVIETETETETETETETETETETETETETETETETETETETETET

```

QY 144 IFISMSVPAVFAFGMIFLELNFGAEIEIYKXHVHCRG---CSVFESKISGVLTFMTSTF 199

142 IGGCWVLPETISELPITMGQNNITGITITREPREKNNNSGATVCTPMNIPDVTTCGGTATF 201

OY 200 YIPGIMLCVYYRIYLAKEQARLISDANOXIIGLEMKNGISQS-----KERKA 249
| | | | : | | | | | | | | | | | | | | | | | |
b 202 YIPFLMIAAYRITYATAKEHNHOT---OMIQRAGASSSSAPDSADONSTDMPTETTKA 257
| | | | : | | | | | | | | | | | | | | | | | |

QY 250 VKTLGIAMGVLLICWCBPFICTVMDPELHYIIPTTLNDVLILFMFGYLNSTFENMVYAEFP 309

258 KTLICIMGCFCOLMAPEFVNIVDEPIDYATPGOWTAPILMIGYTSGLNPTLYAFINLK 317

QY 310 WPKALJOML 319

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2005, 03:30:08 ; Search time 22 Seconds
(without alignments)
1482.613 Million cell updates/sec

Title: US-09-980-145-6
Perfect score: 1806
Sequence: 1 MPPCHNINISCVKNNMSN.....FGKIFQKDSRCKLFLRLSS 339

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	677.5	37.5	2	JC5832	neurotransmitter r
2	529	29.3	387	S55550	5-HT4s receptor -
3	529	29.3	406	S55549	serotonin 4 recept
4	510.5	28.3	459	A56849	dopamine receptor-
5	497.5	27.5	465	I51661	dopamine D1C recep
6	493.5	27.3	363	I50475	dopamine D1 recep
7	486	26.9	358	J01278	histamine H2 recep
8	482	26.7	444	C55886	dopamine receptor
9	482	26.7	450	A55886	dopamine receptor
10	481	26.6	457	I51660	dopamine D1B recep
11	480	26.6	451	I51659	dopamine D1A recep
12	475	26.3	377	S68423	serotonin receptor
13	474	26.2	372	A53279	serotonin receptor
14	473	26.2	374	I77467	histamine H2 recep
15	472.5	26.2	359	JH0449	histamine H2 recep
16	467	25.9	446	I47217	dopamine receptor
17	466.5	25.8	377	B30341	G protein-coupled
18	466.5	25.8	386	S72168	dopamine receptor
19	462.5	25.6	463	B56849	dopamine receptor-
20	461	25.5	483	A25896	beta-adrenergic re
21	459.5	25.4	359	JC4120	histamine H2 recep
22	458	25.4	446	I47217	dopamine receptor
23	454.5	25.2	477	DYH0D5	dopamine receptor
24	452	25.0	418	1 ORHYB2	beta-2-adrenergic
25	451	25.0	418	S10855	beta-2-adrenergic
26	450.5	24.9	359	A39008	histamine H2 recep
27	450	24.9	475	A41271	dopamine receptor
28	449.5	24.9	464	S12591	beta-1-adrenergic
29	449	24.9	418	S00260	beta-2-adrenergic

30	446.5	24.7	415	2	I53040	beta-2 adrenergic
31	446.5	24.7	487	1	DYR7D1	dopamine receptor
32	446	24.7	466	2	S36794	beta-1-adrenergic
33	438.5	24.3	413	1	ORHUB2	beta-2-adrenergic
34	437.5	24.2	486	2	B55886	dopamine receptor
35	435.5	24.0	477	1	QRHUB1	beta-1-adrenergic
36	432.5	23.9	366	2	A47321	serotonin receptor
37	429.5	23.8	428	2	A55044	beta-4C-adrenergic
38	429.5	23.8	477	2	S71323	alpha-1A adrenergic
39	429	23.8	480	2	I53053	beta-1 adrenergic
40	427.5	23.7	366	2	S26048	serotonin receptor
41	427.5	23.7	386	2	A42688	serotonin receptor
42	427.5	23.7	386	2	S18637	serotonin receptor
43	425.5	23.6	429	2	S65656	alpha-1C-adrenergic
44	425.5	23.6	466	2	JN0765	alpha-1C-adrenergic
45	425.5	23.6	499	2	S65657	alpha-1C-adrenergic

ALIGNMENTS

RESULT 1

UC5832
neurotransmitter receptor - human
C/Species: Homo sapiens (man)
C/Date: 20-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004
C/Accession: JC5832
R/Zeng, Z.; Fan, P.; Rand, E.; Kyaw, H.; Su, K.; Madike, V.; Carter, K.C.; Li, Y.
Biochem. Biophys. Res. Commun. 242, 575-578, 1998
A/Title: Cloning of a putative human neurotransmitter receptor expressed in skeletal mus
A/Reference number: JC5832; MUID:98125534; PMID:9464258
A/Accession: JC5832
A/Molecule type: mRNA
A/Residues: 1-337 <ZEN>
A/Cross-references: UNIPROT:O14804; GB:AF021818; NID:92465431; PIDN:AAC39581.1; PID:9246
A/Experimental source: brain
A/Genetics:
A/Map position: 6q23
C/Keywords: glycoprotein; receptor; transmembrane protein
F/29-63/Domain: transmembrane #status predicted <TM1>
F/69-95/Domain: transmembrane #status predicted <TM2>
F/114-118/Domain: transmembrane #status predicted <TM3>
F/149-173/Domain: transmembrane #status predicted <TM4>
F/186-200/Domain: transmembrane #status predicted <TM5>
F/204-229/Domain: transmembrane #status predicted <TM6>
F/253-310/Domain: transmembrane #status predicted <TM7>
F/21/Binding site: carbohydrate (Aan) (covalent) #status predicted

Query Match	37.5%	Score 677.5;	DB 2;	Length 337;
Best Local Similarity	39.0%	Pred. No. 1.3e-49;		
Matches 130;	Conservative 67;	Mismatches 125;	Indels 11;	Gaps 4;
QY	4	FCHNINISCVKNNMSNDVRSALYSIMVLIIITLVGNLIYIVSISHPKQHTPTNMLIH	63	
DB	16	FCYQ-VNGSCRYHTATIGLVYLTCAAGMIIYIGVPAFAVSKALHTPTNFIIL	74	
QY	64	SMATYDFLLGLCINPYSNVAHGWYRGEVPCCKHTSTDMLSASIFHLSFISIDRY	123	
DB	75	SLADLMDPLGLVPLSTIRSVESCFPGDLCRIHTYLDLFCUTSIFHLCFISIDHC	134	
QY	124	AVCDPLRYKAKNNILVICWMIFFISWVPAVFAFGMIFELNPKGABEYVKKVHGSGCS	183	
DB	135	AICDPLLYPSKFTYRVALRYLLAGNGVPAATSLFLYTDVETRLSQ-WLEBMPGVSCQ	193	
QY	184	VFFSKISGVLTFTMSFYIPGSIIMLCVYRYIYLAKEQARLISDANQKLGLEMGKISQ	243	
DB	194	LLINKFWGMNLF-PLFFVPCILIMISLVYKIFVVAATRAQQTTLISKSP-----AGAA	244	
QY	244	SKERAVYITIGIIVGCVFLICWCPFCICVMPPLHYIIPPLINDVLIIFGLINSTFPMV	303	
DB	245	KHRRKAAITIGIIVGVIYLLCWLPTIDYMDVSLHFIPPLVDFIFWFAVFNACNPDI	304	


```
OY 142 VMIISWSYPAVAFQMIIELELNFKGAEIYYKHVHCRGGCSYFSSKIS-GVLTFTMTSPY 200
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 157 TMIAIWMALISICISIPLEFMR-QAKQDEM-----SDCLVNTSQISYTIYSTGARY 207
OY 201 IEGSIMLCVYRIYIAKQO-----AALISDA-----NOKIQIG----- 234
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 208 IPSVLLIILYGRIVAAARRINPRLYGKRFTHALITGSAGSSILCSINSLHEGSHS 267
OY 235 -----LEMKGISQSKERKAVKTLGVMGVFLICMCPEFICTYMDP- 275
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 268 AGSPLEFNNVTKLADSALEKRR-ISAAREKATKILGILGAFITCWLPEFVSLVLP 326
OY 276 -----FLAYIIPPTLNDVLWFGYLNSTNPNVYAFYFWPKALMMLFGKI--FOKDS 328
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 327 CRDSCWIH-----PALDFPFTWLGYLNSLNPILYITVNEBFQA-----FOKIVPRKAS 377
```

```
RESULT 14
177467
serotonin receptor 1D - rat
N.Alternate names: 5-hydroxytryptamine receptor 1D (5-HT1D)
C.Species: Rattus norvegicus (Norway rat)
C.Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #ext_change 09-Jul-2004
C.Accession: I77467
R.Hamblin, M.W.; McGuffin, R.W.; Metcalf, M.A.; Dorsa, D.M.; Merchant, K.M.
Mol. Cell. Neurosci. 3, 578-587, 1992
A.Title: Distinct 5-HT1B and 5-HT1D serotonin receptors in rat: structural and pharmacol
A.Reference number: 157683
A.Accession: I77467
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: DNA
A.Residues: 1-374 <RES>
A.Cross-references: UNIPROT:P28565; GB:M89953; NID:q202544; P1DN:AAA0614.1; P1D:g202545
C.Superfamily: octopamine receptor type I
C.Keywords: neurotransmitter receptor
```

```
Query Match 26.2%; Score 473; DB 2; Length 374;
Best Local Similarity 32.1%; Pred. No. 2.3e-32;
Matches 117; Conservative 62; Mismatches 116; Indels 70; Gaps 10;

OY 18 MSNDV-----RASLSLMLVLIITLVGNLIYVSIHSRQOLHTPTNWLHSMATYDFLLG 73
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 26 WDPEVLQALRIISLVVLSIITLATVLSNAFVLTITLTKKLTTPANVILIGSLATDDLVS 85
OY 74 CLVMPYSWRSAEHCWYGEVFCIKHTSDIMLSASIFPHLSFISIDRYVAVCDPLRYKA 133
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 86 ILVMPISLAVYTTTRTNFGQICDIWSSDITCTASILHLCDVIALDRYMAITDLEYSK 145
OY 134 KNNIIVICWMIFISVPAVAFQMIIELELNFKGAEIYYKHVHCRGGCSYFSSKIS-GV 192
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 146 RRTAGHAAMIAAVALISICISIPLEFMR-QATAHEM-----SDCLVNTSQISYTI 196
OY 193 LTFMTSEYIPGIMLCVYRIYIAKQO-----AALISDA----- 227
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 197 YSTGAFYIPISLIIILYGRIVAAARRINPRLYGKRFTHALITGSAGSSILCSINSLNPS 256
OY 228 -----NQ-----KLQIGLEMKNGISQSKERKAVKTLGVMGVFLICMCPEFI 269
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 257 LHESHTVGSPLPFNNQVKILADSILERKRIISAAREKATKILGILGAFITCWLPEFV 316
OY 270 CTVNDP-----FLHYIIPPTLNDVLWFGYLNSTNPNVYAFYFWPKALMMLFGKI 323
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 317 VSLVPLICRDCWIH-----PALDFPFTWLGYLNSLNPILYITVNEBFQA-----FOKIVPRKAS 377
OY 324 FOKDS 328
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 370 FRKAS 374
```

```
RESULT 15
JH0449
histamine H2 receptor - human
```

```
C.Species: Homo sapiens (man)
C.Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #ext_change 09-Jul-2004
C.Accession: JH0449; I52319
R.Ganz, I.; Munzer, G.; Taehiro, T.; Schaeffer, M.; Wang, L.; DelValle, J.; Yamada, T
Biochem. Biophys. Res. Commun. 178, 1386-1392, 1991
A.Title: Molecular cloning of the human histamine H2 receptor.
A.Reference number: JH0449; MUID:91337087; PMID:1714721
A.Accession: JH0449
```

```
A.Molecule type: DNA
A.Residues: 1-359 <GAN>
A.Cross-references: UNIPROT:P25021; GB:M64799; NID:9184087; P1DN:AAA58647.1; P1D:9184088
R.Nishi, T.; Koike, T.; Oka, T.; Maeda, M.; Futai, M.
Biochem. Biophys. Res. Commun. 210, 616-623, 1995
A.Title: Identification of the promoter region of the human histamine H2-receptor gene.
A.Reference number: I52319; MUID:95275318; PMID:7755641
A.Accession: I52319
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: DNA
A.Residues: 1-359 <RES>
A.Cross-references: GB:P49783; NID:q728495; P1DN:BAA08618.1; P1D:g728496
C.Comment: Histamine is important in the regulation of gastric acid secretion.
C.Genetics:
A.Gene: GDB:HRH2
A.Map position: 5pter-5qter
C.Superfamily: vertebrate rhodopsin
C.Keywords: G protein-coupled receptor; transmembrane protein
F.23-44/Domain: transmembrane #status predicted <TM1>
F.58-81/Domain: transmembrane #status predicted <TM2>
F.93-114/Domain: transmembrane #status predicted <TM3>
F.135-159/Domain: transmembrane #status predicted <TM4>
F.181-204/Domain: transmembrane #status predicted <TM5>
F.235-258/Domain: transmembrane #status predicted <TM6>
F.268-289/Domain: transmembrane #status predicted <TM7>
```

```
Query Match 26.2%; Score 472.5; DB 2; Length 359;
Best Local Similarity 32.6%; Pred. No. 2.5e-32;
Matches 106; Conservative 64; Mismatches 124; Indels 31; Gaps 5;

OY 19 SNDVRAISLMLVLIITLVGNLIYVSIHSRQOLHTPTNWLHSMATYDFLLGCLVMP 78
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 14 STACKITTVLAVLILITVAGNVVCLAVGNRRRLNLTNCFVSLAITDLLGLVLP 73
OY 79 YSWRSAEHCWYGEVFCIKHTSDIMLSASIFPHLSFISIDRYVAVCDPLRYKAKNVL 138
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 74 FSAIYQSLCKRSKPKVCNITYISLDVNLCTASILNLMISLDKCAVMDPLRIPVLPV 133
OY 139 VICWMIFISVPAVAFQMIIELELNFKGAEIYYKHVHCRGGCSYFSSKISGVLTFTMTS 198
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 134 RVALSLVLIWISITLSFSLIHLGMSNRNETS--KNNHTTSCKKVQVNEVYGLVDGLVT 190
OY 199 FYIPGSIIMLCVYRIYIAEQARLISDANQKIQIGLEMKNGISQSK-----ERKAVKT 253
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 191 FYPLPLIMCTIYRIIFVARDQAKRI-----NHSISWGAATIRERKAVTL 236
OY 254 GIWGVFLICMCPEFICTVNA-----DEFLHYIIPPTLNDVLWFGYLNSTNPNVYAFY 308
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 237 AAWVGAFIILCMFPFTFVYRGRLGDADINEV---LEALVNLGYNGLNPILYALN 292
OY 309 PWRPKALMMLFGKIFOKDSRRCKL 333
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 293 RDRFTGYQQLFCRLANRNSHKTSL 317
```

Search completed: February 12, 2005, 03:45:11
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using bw model

Run on: February 12, 2005, 02:35:58 / Search time 72 Seconds
(without alignments)
2411.040 Million cell updates/sec

Title: US-09-980-145-6
Perfect score: 1806
Sequence: 1 MMPCNNINISCVNMSN.....FGXIFQKDSRCKLFLRLSS 339

Scoring table: BLAST62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: uniprot_gprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1806	100.0	339	1	TAR1_HUMAN
2	1757	97.3	338	1	TAR1_MACMU
3	1403	77.7	332	1	TAR1_RAT
4	1351	74.8	332	1	TAR1_MOUSE
5	927.5	51.4	347	1	TAR2_RAT
6	817	45.2	306	2	O6NWS2
7	817	45.2	306	2	O6NWS3
8	815	45.1	306	1	GP58_HUMAN
9	815	45.1	306	2	O6NWS1
10	813.5	45.0	343	1	GP57_HUMAN
11	782	43.3	338	1	TAR3_RAT
12	743	41.1	348	1	TAR3_HUMAN
13	741	41.0	333	1	TAR2_RAT
14	733	40.6	344	1	TAR7_RAT
15	733	40.6	345	1	TAR4_RAT
16	731	40.5	344	1	TAR1_RAT
17	730	40.4	344	1	TAR10_RAT
18	722	40.0	358	1	TAR14_RAT
19	722	40.0	358	1	TAR15_RAT
20	715	39.6	358	1	TAR4_HUMAN
21	713	39.5	345	1	TAR8_RAT
22	710	39.3	358	1	TAR9_RAT
23	704	39.0	358	1	TAR6_RAT
24	697	38.6	342	1	TAR5_HUMAN
25	679	37.6	362	1	TAR3_RAT
26	677.5	37.5	337	2	O14804
27	668.5	37.0	337	2	O6NTH8
28	580.5	32.1	352	2	O9YHV4
29	545.5	30.2	328	2	O9YHV8
30	542.5	30.0	388	1	5H4_CAVPO
31	539	29.8	388	1	5H4_MOUSE

32	534	29.6	388	1	5H4_HUMAN	Q13639	homo sapien
33	534	29.6	388	2	O81XH9	Q81XH9	homo sapien
34	534	29.6	428	2	O712M9	O712M9	homo sapien
35	529	29.3	406	1	5H4_RAT	O62758	rattus norv
36	526.5	29.2	328	2	O9YHV7	O9YHV7	fugu rubrip
37	519	28.7	400	2	O6R1B8	O6R1B8	bus scrofa
38	519	28.7	402	2	O6Q253	O6Q253	bus scrofa
39	511	28.3	353	2	O9YHV3	O9YHV3	lancepeta fl
40	510.5	28.3	459	1	D1DR_FUGRU	P53452	fugu rubrip
41	503	27.9	446	2	O42315	O42315	cyprinus ca
42	498.5	27.6	445	2	O98842	O98842	anguilla an
43	497.5	27.5	465	1	DCDR_XENIA	P42291	xenopus lae
44	495	27.4	437	2	O42316	O42316	cyprinus ca
45	493.5	27.3	363	1	D1DR_CARAU	P35406	carassius a

ALIGNMENTS

RESULT 1
TAR1_HUMAN STANDARD: PRT: 339 AA.
ID TAR1_HUMAN
AC Q96RJO:
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Trace amine receptor 1 (Tar-1).
GN Name=TAR1; Synonyms=TAL, TAR1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Carnivora; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21374364; PubMed=11459929; DOI=10.1073/pnas.151105198;
RA Borowsky B., Adam N., Jones K.A., Raddeatz R., Artyushyn R.,
RA Ogozalek K.L., Durkin M.M., Lakhani P.P., Bonini J.A., Pathirana S.,
RA Boyle N., Pu X., Kouranova E., Lichtblau H., Ochoa F.Y.,
RA Branchek T.A., Gerald C.;
RT "Trace amines: identification of a family of mammalian G protein-
coupled receptors.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:8966-8971(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21580235; PubMed=11723224;
RA Bunzow J.R., Sonders M.S., Artamangkul S., Harrison L.M., Zhang G.,
RA Quigley D.I., Darland T., Suchland K.L., Pasumamula S., Kennedy J.L.,
RA Olsson S.B., Magenis R.E., Amara S.G., Grandy D.K.;
RT "Amphetamine, 3,4-methylenedioxymethamphetamine, lysergic acid
diethylamide, and metabolites of the catecholamine neurotransmitters
are agonists of a rat trace amine receptor.";
RL Mol. Pharmacol. 60:1181-1188(2001).
RN [3]
RP SEQUENCE FROM N.A.
RA Kopatz S.A., Aronstam R.S., Sharma S.V.;
RT "cDNA clones of human proteins involved in signal transduction
stimulated by the G-protein coupled receptor (www.cdna.org).";
RL Submitted (Nov-2002) to the EMBL/Genbank/DBJ databases.
-1- FUNCTION: Orphan receptor. Could be a receptor for trace amines.
Activated by endogenous trace amines as well as metabolites of the
biogenic amine neurotransmitters. Trace amines are biogenic amines
present in very low levels in mammalian tissues. Although some
trace amines have clearly defined roles as neurotransmitters in
invertebrates, the extent to which they function as true
neurotransmitters in vertebrates has remained speculative. Trace
amines are likely to be involved in a variety of physiological
functions that have yet to be fully understood. This receptor
seems to be mediated by the G(s)-class of G-proteins which
activate adenylyl cyclase.
-1- SUBCELLULAR LOCATION: Integral membrane protein.
-1- TISSUE SPECIFICITY: Detected in low levels in discrete regions
within the central nervous system and in several peripheral
tissues. Moderately expressed in stomach. Low levels in amygdala,

CC kidney, and lung, and small intestine. Trace amounts in cerebellum,
CC dorsal root ganglia, hippocampus, hypothalamus, liver, medulla,
CC pancreas, pituitary, pontine reticular formation, prostate,
CC skeletal muscle, and spleen.

CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AF380185; AAK71236.1; -
DR EMBL; AF200627; AAG17112.1; -
DR EMBL; AF180374; AAO22154.1; -
DR Genew; HGNC:17734; TRAR1.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR009132; Tracamine_recept.
DR InterPro; IPR009133; Tracamine_recept1.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PRINTS; PR01831; TRACAMINER.
DR PRINTS; PR01830; TRACAMINER.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Multigene family;
KW Transmembrane.
FT DOMAIN 1 25 Extracellular (Potential).
FT TRANSMEM 26 46 1 (Potential).
FT DOMAIN 47 59 Cytoplasmic (Potential).
FT TRANSMEM 60 80 2 (Potential).
FT DOMAIN 81 98 Extracellular (Potential).
FT TRANSMEM 99 119 3 (Potential).
FT DOMAIN 120 136 Cytoplasmic (Potential).
FT TRANSMEM 137 157 4 (Potential).
FT DOMAIN 158 188 Extracellular (Potential).
FT TRANSMEM 189 209 5 (Potential).
FT DOMAIN 210 252 6 (Potential).
FT TRANSMEM 253 273 Cytoplasmic (Potential).
FT TRANSMEM 274 287 Extracellular (Potential).
FT TRANSMEM 288 308 7 (Potential).
FT DOMAIN 309 339 Cytoplasmic (Potential).
FT DISULFID 36 182 By similarity.
FT CARBOHYD 10 10 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 17 17 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 339 AA; 39091 MW; 5E72FA61CEFA0C00 CRC64;

Query Match 100.0%; Score 1806; DB 1; Length 339;
Best Local Similarity 100.0%; Pred. No. 3.8e-117;
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNPFGHNINISCVKNNNSNDVRASLYSLMWLIITLVGNLIYVISISHPKQALPTNN 60
DB 1 MNPFGHNINISCVKNNNSNDVRASLYSLMWLIITLVGNLIYVISISHPKQALPTNN 60
QY 61 LIHSAATYDFLLGLWMPYSWRSASHEQWGEVCKTHSTDTMLMSASIFHLSTID 120
DB 61 LIHSAATYDFLLGLWMPYSWRSASHEQWGEVCKTHSTDTMLMSASIFHLSTID 120
QY 121 RYVAVCDPLRYKAKNNILVICMIFISWSVAVAFAGMIFELNFKASERYYKHHVCRG 180
DB 121 RYVAVCDPLRYKAKNNILVICMIFISWSVAVAFAGMIFELNFKASERYYKHHVCRG 180
QY 181 GCSVFEKISGVLFPMFSFYIPGSIIMCVYRITLILKEQARLISDANQKLGLEWKG 240
DB 181 GCSVFEKISGVLFPMFSFYIPGSIIMCVYRITLILKEQARLISDANQKLGLEWKG 240
QY 241 ISGSKERKAVLTGLIIVGVPFLICWCPFFICTVMDPEFLHYIIPPLNDVLIWPGYLNSTFN 300
DB 241 ISGSKERKAVLTGLIIVGVPFLICWCPFFICTVMDPEFLHYIIPPLNDVLIWPGYLNSTFN 300

QY 301 PMVYAFEPYFWRKALKMLFGKIFQKDSRRCLFLEISS 339
DB 301 PMVYAFEPYFWRKALKMLFGKIFQKDSRRCLFLEISS 339

RESULT 2

TAR1_MACMU STANDARD; PRT; 338 AA.
AC Q8H264;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Trace amine receptor 1 (Tar-1).
GN Name=TAR1; Synonyms=TAL; TAR1;
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_Taxid=9544;
RN [1]

SEQUENCE FROM N.A.

RA Miller G.W., Madras B.K.;

RT "Cloning of trace amine receptor 1 (TAR1) from Rhesus monkey.";

RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: Orphan receptor. Could be a receptor for trace amines.

CC biogenic amine neurotransmitters (by similarity). Trace amines are

CC although some trace amines have clearly defined roles as

CC neurotransmitters in invertebrates, the extent to which they

CC function as true neurotransmitters in vertebrates has remained

CC speculative. Trace amines are likely to be involved in a variety

CC of physiological functions that have yet to be fully understood.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AY153366; AAN06172.1; -

DR InterPro; IPR000276; GPCR_Rhodopsn.

DR InterPro; IPR009132; Tracamine_recept.

DR InterPro; IPR009133; Tracamine_recept1.

DR Pfam; PF00001; 7tm_1; 1.

DR PRINTS; PR00237; GPCRHHODOPSN.

DR PRINTS; PR01831; TRACAMINER.

DR PRINTS; PR01830; TRACAMINER.

DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.

DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.

KW G-protein coupled receptor; Glycoprotein; Multigene family;
KW Transmembrane.

FT DOMAIN 1 24 Extracellular (Potential).
FT TRANSMEM 25 45 1 (Potential).
FT DOMAIN 46 58 Cytoplasmic (Potential).
FT TRANSMEM 59 79 2 (Potential).
FT DOMAIN 80 97 Extracellular (Potential).
FT TRANSMEM 98 118 3 (Potential).
FT DOMAIN 119 135 Cytoplasmic (Potential).
FT TRANSMEM 136 156 4 (Potential).
FT DOMAIN 157 187 Extracellular (Potential).
FT TRANSMEM 188 208 5 (Potential).
FT DOMAIN 209 251 Cytoplasmic (Potential).
FT TRANSMEM 252 272 6 (Potential).
FT DOMAIN 273 286 Extracellular (Potential).
FT TRANSMEM 287 307 7 (Potential).
FT DOMAIN 308 338 Cytoplasmic (Potential).
FT DISULFID 95 181 By similarity.
FT CARBOHYD 9 9 N-linked (GlcNAc...) (Potential).

FT CARBOHYD 16 16 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 338 AA; 38797 MW; 9EE7B35456B409B CRC64;
 Query Match 97.3%; Score 1757; DB 1; Length 338;
 Best Local Similarity 96.4%; Pred. No. 9,3e-114;
 Matches 326; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 2 MPFCNINISCVKNNKNDVRAISYSLMVLIIITLVGNLIVISISHFKQLHPTTWL 61
 DB 1 MPFCNINISCVKNNKNDVRAISYSLMALIITLVGNLIVISISHFKQLHPTTWL 60
 QY 62 IHSMAVDFLLGCLVMPYSWVSAHCWYFGSEVFCIKHTSTDIMLSASIFHLSPISIDR 121
 DB 61 IHSMAVDFLLGCLVMPYSWVSAHCWYFGSEVFCIKHTSTDIMLSASIFHLSPISIDR 120
 QY 122 YVAVCDPLRYAKKNILVICWIFISWSVPAVAFGMIFLEINFGABEIIYKVVHCRGG 181
 DB 121 YVAVCDPLRYAKKNILVICWIFISWSVPAVAFGMIFLEINFGABEIIYKVVHCRGG 180
 QY 182 CSVFSPKISGVLTFTSTFYIPGSIIMCYVYRIYLAKQARLISDANKLQIGLEMKNGI 241
 DB 181 CSVFSPKISGVLTFTSTFYIPGSIIMCYVYRIYLAKQARLISDANKLQIGLEMKNGI 240
 QY 242 SOSKERKAVKTLGIWGVFLICWCPFICTVMDPELHYIIPPTLNDVLIWFGYLNSTNP 301
 DB 241 SOSKERKAVKTLGIWGVFLICWCPFICTVMDPELHYIIPPTLNDVLIWFGYLNSTNP 300
 QY 302 MYVAFYPMFRKALFMFLFGKIFORDSSRCKLFLELSS 339
 DB 301 MYVAFYPMFRKALFMFLFGKIFORDSSRCKLFLELSS 338

RESULT 3
 TART1_RAT STANDARD; PRT; 332 AA.
 ID TART1_RAT
 AC 0923Y9; Q8VH05;
 DT 29-MAR-2004 (Rel. 43; Created)
 DT 29-MAR-2004 (Rel. 43; Last sequence update)
 DT 25-OCT-2004 (Rel. 45; Last annotation update)
 DE Trace amine receptor 1 (Tart-1).
 GN Name=Tart1; Synonym=Ta1, TART1;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CX NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANT GLN-170.
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=21374364; PubMed=11459929; DOI=10.1073/pnas.151105198;
 RA Borowsky B., Adham N., Jones K.A., Raddatz R., Artymshyn R.,
 RA Ogorzalek K.L., Durkin M.W., Lakhani P.P., Bonini J.A., Pathirana S.,
 RA Boyle N., Pu X., Kouranova E., Lichtblau H., Ochoa F.Y.,
 RA Branchek T.A., Gerald C.;
 RT "Trace amines: identification of a family of mammalian G protein-
 RT coupled receptors.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:8966-8971(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Cerebellum, and Pancreatic tumor;
 RX MEDLINE=21580235; PubMed=11723224;
 RA Bunzow J.R., Sonders M.S., Artamangkul S., Harrison L.M., Zhang G.,
 RA Oulley D.I., Darland T., Suchland K.L., Pasumamula S., Kennedy J.L.,
 RA Olson S.B., Magenis R.E., Amara S.G., Grandy D.K.;
 RT "Amphetamine, 3,4-methylenedioxymethamphetamine, lysergic acid
 RT diethylamide, and metabolites of the catecholamine neurotransmitters
 RT are agonists of a rat trace amine receptor.";
 RL Mol. Pharmacol. 60:1181-1188(2001).
 CC -1- FUNCTION: Orphan receptor. Could be a receptor for trace amines.
 CC Activated by endogenous trace amines as well as metabolites of the
 CC biogenic amine neurotransmitters. Trace amines are biogenic amines
 CC present in very low levels in mammalian tissues. Although some
 CC trace amines have clearly defined roles as neurotransmitters in
 CC invertebrates, the extent to which they function as true

CC neurotransmitters in vertebrates has remained speculative. Trace
 CC amines are likely to be involved in a variety of physiological
 CC functions that have yet to be fully understood. This receptor
 CC seems to be mediated by the G(s)-class of G-proteins which
 CC activate adenylylate cyclase.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Widely distributed, but in low abundance,
 CC throughout the brain. Highest levels detected in the olfactory
 CC bulb, nucleus accumbens/olfactory tubercle, prefrontal cortex and
 CC other cortical regions, midbrain regions consisting of substantia
 CC nigra and ventral tegmentum, cerebellum, and pons/medulla. Among
 CC peripheral tissues, highest level observed in the liver, less
 CC expression in kidney, gastrointestinal tract, spleen, pancreas,
 CC and heart.
 CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF380186; AAK71237.1; -;
 CC EMBL; AF421352; AAL65137.1; -;
 CC RGD; 621621; Tart1.
 CC InterPro; IPR000276; GPCR_Rhodopsn.
 CC InterPro; IPR009132; Tracamine_recept.
 CC InterPro; IPR009133; Tracamine_receptl.
 CC Pfam; PF00001; 7tm_1; 1.
 CC PRINTS; PR00237; GPCR_HODOPSIN.
 CC PRINTS; PR01831; TRACAMINER.
 CC PRINTS; PR01830; TRACAMINER.
 CC PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
 CC PROSITE; PS50262; G_PROTEIN_RECP_F1_2; 1.
 CC G-protein coupled receptor; Glycoprotein; Multigene family;
 CC Polymorphism; Transmembrane.
 CC DOMAIN 1 24
 CC TRANSMEM 25 45
 CC FT TRANSMEM 46 58
 CC FT DOMAIN 59 79
 CC FT TRANSMEM 80 97
 CC FT TRANSMEM 98 118
 CC FT DOMAIN 119 138
 CC FT TRANSMEM 139 159
 CC FT DOMAIN 160 187
 CC FT TRANSMEM 188 208
 CC FT DOMAIN 209 249
 CC FT TRANSMEM 250 270
 CC FT DOMAIN 271 287
 CC FT TRANSMEM 288 308
 CC FT DOMAIN 309 332
 CC FT DISULFID 95 181
 CC FT CARBOHYD 9 16
 CC FT CARBOHYD 16 16
 CC FT CARBOHYD 283 283
 CC FT VARIANT 170 170
 CC SQ SEQUENCE 332 AA; 38021 MW; AD7F3A728C77B246 CRC64;
 Query Match 77.7%; Score 1403; DB 1; Length 332;
 Best Local Similarity 78.1%; Pred. No. 2,7e-89;
 Matches 261; Conservative 27; Mismatches 44; Indels 2; Gaps 1;

QY 2 MPFCNINISCVKNNKNDVRAISYSLMVLIIITLVGNLIVISISHFKQLHPTTWL 61
 DB 1 MPFCNINISCVKNNKNDVRAISYSLMALIITLVGNLIVISISHFKQLHPTTWL 60
 QY 62 IHSMAVDFLLGCLVMPYSWVSAHCWYFGSEVFCIKHTSTDIMLSASIFHLSPISIDR 121
 DB 61 IHSMAVDFLLGCLVMPYSWVSAHCWYFGSEVFCIKHTSTDIMLSASIFHLSPISIDR 120
 QY 122 YVAVCDPLRYAKKNILVICWIFISWSVPAVAFGMIFLEINFGABEIIYKVVHCRGG 181

```

Cc 121 YVAVCDPLRYKAKINLALFVMIILMSLPAVAFEGMFLFELNLSGVEELYNQVCLRG 180
Cc 182 CSVFESKISGVLTFMTSTFYIPGSMLCVYRYIYIAEQARLISDANOKLQIGLEMNGI 241
Cc 181 CFFPFSKVSGLVAMTSTFYIPGSMVLFVYRYIYIAKQARSINRAN--LVQGLEGSRA 238
Cc 242 SOSKERKAVKTLGIWGVFLICWCPFFICTVMDPELHYIIPPTLNDVLIWFGYLNSTNP 301
Cc 239 POSKETAKTLGIWGVFLICWCPFFICTVMDPELHYIIPPTLNDVLIWFGYLNSTNP 298
Cc 302 MYAFYFPWFRKALKMLFGKIFQKDSRSCKLFL 335
Cc 299 MYAFYFPWFRKALKMLFGKIFQKDSRSCKLFL 332

RESULT 4
ID TAR1_MOUSE STANDARD; PRT; 332 AA.
AC Q92378;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Trace amine receptor 1 (Tar1).
GN Name-Tar1; Synonyms=Tal, Tar1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVU;
RX MEDLINE=21374364; PubMed=11459929; DOI=10.1073/pnas.151105198;
RA Borowsky B., Adham N., Jones K.A., Raddatz R., Artymushyn R.,
RA Ogozalek K.L., Durkin M.M., Lakhiani P.P., Bonini J.A., Pathirana S.,
RA Boyle N., Pu X., Kouranova E., Lichtblau H., Octoa F.Y.,
RA Branchek T.A., Gerald C.;
"Trace amines: identification of a family of mammalian G protein-
coupled receptors."
Proc. Natl. Acad. Sci. U.S.A. 98:8966-8971(2001).
-1- FUNCTION: Orphan receptor. Could be a receptor for trace amine.
Activated by endogenous trace amines as well as metabolites of the
biogenic amine neurotransmitters (By similarity). Trace amines are
biogenic amines present in very low levels in mammalian tissues.
Although some trace amines have clearly defined roles as
neurotransmitters in invertebrates, the extent to which they
function as true neurotransmitters in vertebrates has remained
speculative. Trace amines are likely to be involved in a variety
of physiological functions that have yet to be fully understood.
This receptor seems to be mediated by the G(s)-class of G-proteins
which activate adenylate cyclase (By similarity).
-1- SUBCELLULAR LOCATION: Integral membrane protein.
-1- TISSUE SPECIFICITY: Widely distributed throughout the brain.
Strongly expressed in the mitral cell layer of the olfactory bulb,
piriform cortex, the arcuate, motor, and mesencephalic trigeminal
nuclei, lateral reticular and hypoglossal nuclei, cerebellar
Purkinje cells, and ventral horn of the spinal cord. Moderately
expressed in the frontal, entorhinal, and agranular cortices, the
ventral pallidum, thalamus, hippocampus, several hypothalamic
nuclei, amygdala, dorsal raphe, and gigantocellular reticular
nuclei. Weakly expressed in the septum, basal ganglia, amygdala,
mesencephalon, and spinal cord dorsal horn. Particularly
interesting is the moderate expression in several monoaminergic
cell groups, namely the dorsal raphe, the locus coeruleus, and the
ventral tegmental area.
-1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
-1- SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.ebi.ac.uk/announcements/

```

```

Cc or send an email to license@sib.ch).
Cc EMBL; AF380187; AK71238.1; -.
Cc MGD; MG1:2148258; Tar1.
Cc GO; GO:0015021; C:integral to membrane, IC.
Cc InterPro; IPR000276; GPCR_Rhodopsn.
Cc InterPro; IPR009132; Tracamine_recept.
Cc InterPro; IPR009133; Tracamine_recept1.
Cc Pfam; PF00001; 7tm_1; 1.
Cc DR PRINTS; PR00237; GPCR_RHODOPSIN.
Cc DR PRINTS; PR01831; TRACAMINER.
Cc DR PRINTS; PR01830; TRACAMINER.
Cc DR PROSITE; PS00237; G-PROTEIN RECEPTOR F1.1; 1.
Cc DR PROSITE; PS50262; G-PROTEIN RECEPTOR F1.2; 1.
Cc KW G-protein coupled receptor; Glycoprotein; Multigene family;
Cc Transmembrane.
Cc FT DOMAIN 1 24 Extracellular (Potential).
Cc FT TRANSMEM 25 45 1 (Potential).
Cc FT DOMAIN 46 58 Cytoplasmic (Potential).
Cc FT TRANSMEM 59 79 2 (Potential).
Cc FT DOMAIN 80 97 Extracellular (Potential).
Cc FT TRANSMEM 98 118 3 (Potential).
Cc FT DOMAIN 119 138 Cytoplasmic (Potential).
Cc FT TRANSMEM 139 159 4 (Potential).
Cc FT DOMAIN 160 187 Extracellular (Potential).
Cc FT TRANSMEM 188 208 5 (Potential).
Cc FT DOMAIN 209 249 Cytoplasmic (Potential).
Cc FT TRANSMEM 250 270 6 (Potential).
Cc FT DOMAIN 271 287 Extracellular (Potential).
Cc FT TRANSMEM 288 308 7 (Potential).
Cc FT DOMAIN 309 332 Cytoplasmic (Potential).
Cc FT DISULFID 95 181 By similarity.
Cc FT CARBOHYD 9 N-linked (GlcNAc...) (Potential).
Cc SQ SEQUENCE 332 AA; 37620 MW; 680B692B1892264 CRC64;

Query Match 74.8%; Score 1351; DB 1; Length 332;
Best Local Similarity 75.4%; Pred. No. 1e-85;
Matches 252; Conservative 31; Mismatches 49; Indels 2; Gaps 1;

Cc 2 MPFENININISCVKNNMSNDVRSALYSMLVLIITLVNGTLVIYVSIHFKQLHTPTNML 61
Cc 1 MHCHALTNINSHNSWMSREVQASLYSLMSLITLITLVNGTLVIYVSIHFKQLHTPTNML 60
Cc IHSMAVDFPLGLVLPYNSVWVSAHECWYEGVEFCIKHSTDMSSASIFHLSFISIDR 121
Cc 61 LHSMAVDFPLGLVLPYNSVWVSAHECWYEGVEFCIKHSTDMSSASIFHLSFISIDR 120
Cc 122 YVAVCDPLRYKAKINLALFVMIILMSLPAVAFEGMFLFELNLSGVEELYNQVCLRG 181
Cc 121 YVAVCDPLRYKAKINLALFVMIILMSLPAVAFEGMFLFELNLSGVEELYNQVCLRG 180
Cc 182 CSVFESKISGVLTFMTSTFYIPGSMLCVYRYIYIAEQARLISDANOKLQIGLEMNGI 241
Cc 181 CFFPFSKVSGLVAMTSTFYIPGSMVLFVYRYIYIAKQARSINRAN--LVQGLEGSRA 238
Cc 242 SOSKERKAVKTLGIWGVFLICWCPFFICTVMDPELHYIIPPTLNDVLIWFGYLNSTNP 301
Cc 239 POSKETAKTLGIWGVFLICWCPFFICTVMDPELHYIIPPTLNDVLIWFGYLNSTNP 298
Cc 302 MYAFYFPWFRKALKMLFGKIFQKDSRSCKLFL 335
Cc 299 MYAFYFPWFRKALKMLFGKIFQKDSRSCKLFL 332

RESULT 5
ID TAR2_RAT STANDARD; PRT; 347 AA.
AC Q92377;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Trace amine receptor 2 (Tar2).
GN Name-Tar2; Synonyms=Taz, Tar2;

```

Df		14	FCEPAAANNCSPPKAPALVACAMYLVMGAIWMTLMGMVVIIISIAHFQKHSPNLFIL	73
Oy		64	SNAIYDPRLLGLVMPYSWMRASEHCWGCVFCKIHSTDIIMLSAISIPIHLSIDIRY	123
Df		74	SNATYDFLLSCVMPPEPSMWRSIESCWYGDLFCRKHSCCDMLCTTTSIPHLCFISVRHY	133
Oy		124	AACDPFLRYAKNNIVICWVFIFSWVPVAFAFGMIFLELFPKAEEIYGHGVCRSGCS	183
Df		134	AVCDPLHYTOTTTGVGVGYLLINSVYILFRFGLVSESLNLIGED-FVAALICTGLCV	197
Oy		184	VFEKSISGVLTREMTSFYIPGSIMLCVYRIYILAKEOARLISDNQKLQIGLEMKNISQ	243
Df		193	LIEFKLMVLASFIAPFLPGALMGVIYIHFVARKHARKIGCPRTRKALSESMMATS	255
Oy		244	SKERKAVKTLGIYMGVFLCNCPEFFICVMDPPLHYITLPFLNDVLMFGVLTNFEMPV	303
Df		253	GKESKATKTLSIVMGFWCLCMPFVLTITDPFGFTTPEDLYNVFLMLGYNSTFNPII	312
Oy		304	YAFEPMPFKALKOMLFGIKFOKSDSRCL	333
Df		313	YCMFYPMFKRLMYVTGITTFRSDBSSTSLSL	342
RESULT 6				
O6NMS2				
ID	O6NMS2	PRELIMINARY;	PRT;	306 AA.
AC	O6NMS2			
DT	05-JUL-2004 (TrEMBLrel. 27, Created)			
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)			
DT	05-JUL-2004 (TrEMBLrel. 27, Laet annotation update)			
DE	GP858 protein (Fragment).			
CN	Name=GP858;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=PCR rescued clones;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klauner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhac N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang Y., Hasten F.,			
RA	Diatchenko L., Marcusna K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Uedon T.B., Toshiyuki S., Carinici P., Prange C.,			
RA	Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,			
RA	Bohak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.U., Huliy S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Faley J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schultz J., Myers R.M., Butterfield Y.S.,			
RA	Krzywnicki M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,			
RA	Jones S.J., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences."			
RU	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=PCR rescued clones;			
RA	Strauberg R.;			
RL	Submitted (MAR-2004) to the EMBL/Genbank/DBJ databases.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).			
CC	-1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.			
DR	EMBL: BC067462; AA067462.1; -			
DR	GO: GO:0016021; C:integral to membrane; IEA.			
DR	GO: GO:0004872; C:receptor activity; IEA.			
DR	GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.			
DR	GO: GO:0007186; P:G-protein coupled receptor protein signaln.; IEA.			
DR	Interpro: IPR000276; GPCK_Rhodopsn			

```

DR   InterPro: IPR009132; Tracamine_recept.
DR   Pfam: PF00001; 7tm_1, 1.
DR   PRINTS: PR00237; GPCRHOODPSN.
DR   PRINTS: PR01830; TRACEAMINER.
DR   PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
FT   NON_TER 306
SQ   SEQUENCE 306 AA; 34938 MW; 82935D280DDE3C6F CRC64;

Query Match      45.2%; Score 817; DB 2; Length 306;
Best Local Similarity 50.0%; Pred. No. 7,8e-49;
Matches 158; Conservative 49; Mismatches 85; Indels 24; Gaps 4;

QY  26 LYSIMVLIIITLVGNLIVIVSISHPKQHTPTNMLIHSNATVDFLGLCLVMPYSWVRS 85
DB  1 MYSFMAGSIPTITFGNLMIIISIFKQHTPTNFIILSMATIDFLGLFTIMPYSMIRSV 60
QY  86 EHCWYFGEVPCFKHTSDIMLSASIFHLISIDRYAVCDPLRYKAKNIIIVICWTF 145
DB  61 ENCWYFGLTCKTIYSFDMLSTISIFHLCSVAIDRFVACVPLVSTKITTPVIRKLL 120
QY  146 ISMSVPAVFAFGMIFELNFKGAEIYYKHVCRGCGSVFSSKISGVLTFMTSFYIPGSI 205
DB  121 LCMWVPGAFAGVGFSEAVADIG- YDILVACSSSCPWFNKLWGTTLFPMAGFTPGSM 179
QY  206 MLCVYRIYLIAGEQRLISD--ANOKLOGLEMKNGISQSKERKAVKTIGIYGVFLIC 263
DB  180 MGVIGIKIFAVSRKHAIINLNRENQNV-----KKDKAKAKTIGIYGVFLIC 229
QY  264 WCPFFICTVMDPFLHAIPTLNDVLIWPGYLNSTENPMWYAFYFWPFRKALQMFLFGKI 323
DB  230 WFPCCFTIILDPLNLTSTVPLFDALTWFGYFNSTONPLIYGFYFWFRALKYVLLGKI 289
QY  324 F-----QKDS 328
DB  290 FSSCFHNTILCMQKES 305

RESULT 7
ID   Q6NWS3; PRELIMINARY; PRT; 306 AA.
AC   Q6NWS3;
DT   05-JUL-2004 (Tremblrel. 27, Created)
DT   05-JUL-2004 (Tremblrel. 27, Last sequence update)
DE   05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE   GPR58 protein (Fragment).
GN   Name=GPR58;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX   NCBI_Taxid=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PCR rescued clones;
RX   MEDLINE=2238257; PubMed=1477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner K.H., Scheiner C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,
RA   Ditchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stappleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,
RA   Villalón D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Heitlen E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA   Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmitt J., Myers R.M., Buterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smallus D.E., Schmeich A., Schein J.E.,
RA   Jones S.J., Maitra M.A.;

```

```

RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences."
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN   [2]
RC   SEQUENCE FROM N.A.
RP   TISSUE=PCR rescued clones;
RA   Strausberg R.;
RL   Submitted (MAR-2004) to the EMBL/Genbank/DBSJ databases.
CC   - SUBCELLULAR LOCATION: Integral membrane protein (by similarity).
CC   - SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
DR   EMBL: BC067461; AAH67461.1; -.
DR   GO: GO:0016021; C:integral to membrane; IEA.
DR   GO: GO:0004872; F:receptor activity; IEA.
DR   GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO: GO:0007186; P:G-protein-coupled receptor protein signaling. . . IEA.
DR   InterPro: IPR00276; GPCR_Rhodpsn.
DR   InterPro: IPR009132; Tracamine_recept.
DR   Pfam: PF00001; 7tm_1, 1.
DR   PRINTS: PR00237; GPCRHOODPSN.
DR   PRINTS: PR01830; TRACEAMINER.
DR   PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
FT   NON_TER 306
SQ   SEQUENCE 306 AA; 34952 MW; 2D935D280DDE2A03 CRC64;

Query Match      45.2%; Score 817; DB 2; Length 306;
Best Local Similarity 50.0%; Pred. No. 7,8e-49;
Matches 158; Conservative 49; Mismatches 85; Indels 24; Gaps 4;

QY  26 LYSIMVLIIITLVGNLIVIVSISHPKQHTPTNMLIHSNATVDFLGLCLVMPYSWVRS 85
DB  1 MYSFMAGSIPTITFGNLMIIISIFKQHTPTNFIILSMATIDFLGLFTIMPYSMIRSV 60
QY  86 EHCWYFGEVPCFKHTSDIMLSASIFHLISIDRYAVCDPLRYKAKNIIIVICWTF 145
DB  61 ENCWYFGLTCKTIYSFDMLSTISIFHLCSVAIDRFVACVPLVSTKITTPVIRKLL 120
QY  146 ISMSVPAVFAFGMIFELNFKGAEIYYKHVCRGCGSVFSSKISGVLTFMTSFYIPGSI 205
DB  121 LCMWVPGAFAGVGFSEAVADIG- YDILVACSSSCPWFNKLWGTTLFPMAGFTPGSM 179
QY  206 MLCVYRIYLIAGEQRLISD--ANOKLOGLEMKNGISQSKERKAVKTIGIYGVFLIC 263
DB  180 MGVIGIKIFAVSRKHAIINLNRENQNV-----KKDKAKAKTIGIYGVFLIC 229
QY  264 WCPFFICTVMDPFLHAIPTLNDVLIWPGYLNSTENPMWYAFYFWPFRKALQMFLFGKI 323
DB  230 WFPCCFTIILDPLNLTSTVPLFDALTWFGYFNSTONPLIYGFYFWFRALKYVLLGKI 289
QY  324 F-----QKDS 328
DB  290 FSSCFHNTILCMQKES 305

RESULT 8
ID   GP58_HUMAN STANDARD; PRT; 306 AA.
AC   GP58;
DT   29-MAR-2004 (Rel. 43, Created)
DT   29-MAR-2004 (Rel. 43, Last sequence update)
DE   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Probable G protein-coupled receptor GPR58.
GN   Name=GPR58;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX   NCBI_Taxid=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20149852; PubMed=10684976; DOI=10.1016/S0167-4781(99)00241-9;
RA   Lee D.K., Lynch K.R., Nguyen T., Im D.-S., Cheng R., Saldivia V.R.,
RA   Liu Y., Liu I.S.C., Heng H.H.Q., Seeman P., George S.R., O'Dowd B.F.,

```


RA Marchese A.; Cloning and characterization of additional members of the G protein-coupled receptor family.";
 RT Biochim. Biophys. Acta 1490:311-323(2000).
 RU -1- FUNCTION: Orphan receptor.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Not expressed in the pons, thalamus, hypothalamus, hippocampus, caudate, putamen, frontal cortex, basal forebrain, midbrain or liver.
 CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF112460; AAF27278.1; -.
 CC Gene: HGNC:4514; GPR58.
 DR MIM: 604849; -.
 DR GO: GO:0004930; F-G-protein coupled receptor activity; TAS.
 DR InterPro: IPR000237; GPCR_Rhodopsn.
 DR InterPro: IPR009132; Tracamine_recept.
 DR Pfam: PF00001; 7tm1; 1.
 DR PRINTS: PR00237; GPCR_RHODOPSIN.
 DR PRINTS: PR01830; TRACAMINER.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Glycoprotein; Transmembrane. Extracellular (Potential).
 FT DOMAIN 1 3
 FT TRANSMEM 4 24
 FT TRANSMEM 25 34
 FT TRANSMEM 35 55
 FT TRANSMEM 56 73
 FT TRANSMEM 74 94
 FT TRANSMEM 95 117
 FT TRANSMEM 118 138
 FT TRANSMEM 139 162
 FT TRANSMEM 163 183
 FT TRANSMEM 184 218
 FT TRANSMEM 219 239
 FT TRANSMEM 240 254
 FT TRANSMEM 255 277
 FT TRANSMEM 278 306
 FT DOMAIN 71 156
 FT DISULFID 244 244
 FT CARBOHYD 244 244
 SQ SEQUENCE 306 AA; 34924 MW; 55629F613062777C CRC64;
 Query Match 45.1%; Score 815; DB 1; Length 306;
 Best Local Similarity 50.0%; Pred. No. 1.1e-48;
 Matches 158; Conservative 48; Mismatches 86; Indels 24; Gaps 4;
 QY 26 LYSIMLTLLITLVGNLIVIVSISHFQKLTHTPTNMLHISMATVDLLGLCLVMPYSMTASA 85
 DB 1 MYSPFAGSIFITIFGNLMIITISISYFKQLHTPTNLLISMATDPLGLGFTIMPYSMTSV 60
 QY 86 EHCWFGVGFCKIHTSTDIMLSSASIFHLSPISIDRYAVCDPLRYKAMNLTVCWIF 145
 DB 61 ENCWFGTLPFCIKTYSPFLMSTITIFHLCSVALIDRFALICPLLYSTKITITIPVKRLIL 120
 QY 146 ISMSVPAVAFARGMIFLELNFKAEEIYKWHVCRGGCVFPGKISGVLTPMSPFIIPSI 205
 DB 121 LCMSPVGAFAFGAVSEAVADIEG-YDILVACSSCPFMWKKLGTLLFMAGFTPSM 179
 QY 206 MLCVYRYIYLAKEQARLISD--ANQKQIGLEMKNGISQSKERKAVTTLGIWGVFLIC 263
 DB 180 MVGIGIKLFAVSRKAHAINNRQNNQV-----KDKKAAKTLGIVIGVFLIC 229
 QY 264 WCPPICTVMDPFLHYIITPTLNDVLIWFGYNSTFNPMTVAFFTPWRKALKQMLFGKI 323
 DB 230 WPCPFITLLDPFLNFTPVVLFDALTWFGYNSTCNPLIYGFYPWFRALKYILGKI 289

QY 324 F-----QKDS 328
 DB 290 FSSCFNHTILCMOKES 305
 RESULT 9
 Q6NWS1 PRELIMINARY; PRT; 306 AA.
 AC Q6NWS1;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DE GPR58 protein (Fragment).
 GN Name=GPR58;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PCR rescued clones;
 RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Ditschenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Ueda T.B., Toshimuki S., Carrinci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalls D.E., Schermer A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PCR rescued clones;
 RA Strausberg R.;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 DR EMBL: BC067463; AA067463.1; -.
 DR GO: GO:0016021; C:Integral to membrane; IEA.
 DR GO: GO:0004872; F:receptor activity; IEA.
 DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO: GO:0007186; P:G-protein coupled receptor protein signaln. .; IEA.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR009132; Tracamine_recept.
 DR Pfam: PF00001; 7tm1; 1.
 DR PRINTS: PR00237; GPCR_RHODOPSIN.
 DR PRINTS: PR01830; TRACAMINER.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Receptor; Transmembrane.
 FT NON_TER 306 306
 SQ SEQUENCE 306 AA; 34979 MW; 2D93423607FFA03 CRC64;
 Query Match 45.1%; Score 815; DB 2; Length 306;
 Best Local Similarity 50.0%; Pred. No. 1.1e-48;
 Matches 158; Conservative 49; Mismatches 85; Indels 24; Gaps 4;
 QY 26 LYSIMLTLLITLVGNLIVIVSISHFQKLTHTPTNMLHISMATVDPLGLCLVMPYSMTASA 85
 DB 1 MYSPFAGSIFITIFGNLMIITISISYFKQLHTPTNLLISMATDPLGLGFTIMPYSMTSV 60

```

QY 86 EHCWYFGEVCKIHTSTDMILSSASIFHLSPISIDRYAVCDPLRYKAKNIIIVICMIF 145
Db 61 ENCWYFGLTFCCKIYSPDMLSTISIFHLCSVAIDRYALCYPLISTKTIITIVIKRL 120
QY 146 ISMSVAVPAFAGMIFELNKRGAETIYKAVHCRGGSVFPSKISGVLTMTSPYIPGSI 205
Db 121 LCVSWPAGFAFAGVVFSEAVADGIEG-YDILVACSSSCPVMFNKLTGTLTFMAGFTPGSM 179
QY 206 MLCVYRIVYLAKEOARLISD--ANOKLOIGLEMKNGISOSKERKAVKTIGIMGVFLIC 263
Db 180 MGVYKGIYFAVSRKHAIAINNRKNQV-----KKQKKAKTIGIVGVFLIC 229
QY 264 WCPFFITVMDPLHYIIPPLNDVLIWFGYLNSTPNVYAFYFPMFRALKMLFGKI 323
Db 230 WPCFFITLIDPLFNSTPVVLFDAIKMFGYFNSTCNPLIGYFFYPMFRALKYIILGKI 269
QY 324 F-----OKDS 328
Db 290 FSSCFNHTIICWQES 305

```

RESULT 10

GP57_HUMAN STANDARD; PRT; 343 AA.

```

AC Q9P1F4;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Probable G protein-coupled receptor GPR57.
GN Name=GPR57;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20149852; PubMed=10684976; DOI=10.1016/S0167-4781(99)00241-9;
RA Lee D.K., Lynch K.R., Nguyen T., Im D.-S., Cheng R., Saldivia V.R.,
RA Liu Y., Liu I.S.C., Heng H.H.Q., Seeman P., George S.R., O'Dowd B.F.,
RA Marchese A.;
RT "Cloning and characterization of additional members of the G protein-
RT coupled receptor family."
RL Biochim. Biophys. Acta 1490:311-323 (2000).
CC -!- FUNCTION: Orphan receptor.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Not expressed in the pons, thalamus, globus
CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF112461; AAF27279.1; -.
DR GenBank; HGNC:4513; GPR57.
DR MIM; 604848; -.
DR GO; GO:0004930; F-G-protein coupled receptor activity; TAS.
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR InterPro; IPR009132; Tracamine_recept.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PRO0237; GPCR_RHODOPSIN.
DR PROSITE; PRO1630; TRACEAMINER.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Transmembrane.
FT DOMAIN 1 35 Extracellular (Potential).
FT TRANSMEM 36 56 Transmembrane (Potential).
FT DOMAIN 57 68 Cytoplasmic (Potential).

```

```

FT TRANSMEM 69 89 2 (Potential).
FT DOMAIN 90 150 Extracellular (Potential).
FT TRANSMEM 151 168 3 (Potential).
FT DOMAIN 169 172 Cytoplasmic (Potential).
FT TRANSMEM 173 193 4 (Potential).
FT DOMAIN 194 198 Extracellular (Potential).
FT TRANSMEM 199 223 5 (Potential).
FT DOMAIN 224 257 Cytoplasmic (Potential).
FT TRANSMEM 258 278 6 (Potential).
FT DOMAIN 279 287 Extracellular (Potential).
FT TRANSMEM 288 308 7 (Potential).
FT DOMAIN 309 343 Cytoplasmic (Potential).
FT CAROXYD 145 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 343 AA; 39065 MW; DC51D5ABBC026F CRC64;

```

Query Match 45.0%; Score 813.5; DB 1; Length 343;
 Best Local Similarity 46.0%; Pred. No. 1.5e-48;
 Matches 159; Conservative 57; Mismatches 99; Indels 31; Gaps 5;

```

QY 5 CHNIIN-----ISGVKN-----WSNDVRSLSXSNVLIITLVGNLIYVSGIS 50
Db 13 CPKFNKILSSHOPLFSCPDNVFQYDMSHDY-----PLFGLNLIWVSGIS 58
QY 51 FKQHTPTNMLIHSMAVDPLGLAVMPYSWRSABHCWYFGEVCKIHTSTDMILSSAS 110
Db 59 FKQHSPTNMLISMATTDPLGLVIMPYISMSVSCWIFGDFCKFHTSPDMRLTS 118
QY 111 IFHLSPISIDRYAVCDPLRYKAKNIIIVICMIFISWVPAVAFGMIFLEINFGKAE 170
Db 119 IFHLCSIAIDRFYAVCYPLHYTMTNSTIKQLAFWSPALFSPFLVSEADVSGMQS 178
QY 171 IYGHVHCRGGSVFPSKISGVLTMTSPFIRPSIMLCVYRYLYLAKEARLISDANCK 230
Db 179 YKLIIVACNFCALTFKFKGTILFTCFPTPSIMWIGIKLPIYSKOHARVISHPEN 237
QY 231 LQIGLEMKNGISOSKERKAVKTIGIMGVFLICWCEFFICTVMDPLHYIIPPLNDVLI 290
Db 238 TKGK--VKRLSKKORAKKTIGIWMGVFLACMLDCLFVLAVLDPLVDISTPILIIDLV 295
QY 291 WFGYLNSTPNVYAFYFPMFRALKMLFGKIFQDSSRCKFLR 336
Db 296 WLRVFNSTCNPLIHGFENPWFQAFYIVSGKIFSSHSETANLPE 341

```

RESULT 11

TAR3_RAT STANDARD; PRT; 338 AA.

```

AC Q923T6;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Trace amine receptor 3 (Tar-3).
GN Name=Trar3; Synonyms=Tr3; Tar3;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=21374364; PubMed=11459929; DOI=10.1073/pnas.151105198;
RA Borowsky B., Adham N., Jones K.A., Raddatz R., Artymushyn R.,
RA Gozales K.L., Durkin M.M., Lakhani P.P., Bonini J.A., Pathirana S.,
RA Boyle N., Pu X., Kouranova E., Lichtblau H., Ochoa F.Y.,
RA Branchek T.A., Gerald C.;
RT "Trace amines: identification of a family of mammalian G protein-
RT coupled receptors."
RL Proc. Natl. Acad. Sci. U.S.A. 98:8966-8971 (2001).
CC -!- FUNCTION: Orphan receptor. Could be a receptor for trace amines.
CC Trace amines are biogenic amines present in very low levels in
CC mammalian tissues. Although some trace amines have clearly defined
CC roles as neurotransmitters in invertebrates, the extent to which
CC they function as true neurotransmitters in vertebrates has

```


FT DOMAIN 169 197 Extracellular (Potential).
FT TRAMSEM 198 218 5 (Potential).
FT DOMAIN 219 259 Cytoplasmic (Potential).
FT TRAMSEM 260 280 6 (Potential).
FT DOMAIN 281 294 Extracellular (Potential).
FT TRAMSEM 295 315 7 (Potential).
FT DOMAIN 316 348 Cytoplasmic (Potential).
FT DISULFID 105 190 By similarity.
FT CARBOHYD 4 4 N-linked (GlcNAc..?) (Potential).
FT CARBOHYD 19 19 N-linked (GlcNAc..?) (Potential).
SQ SEQUENCE 348 AA; 39016 MW; D0C648480A6F761 CRC64;

Query Match 41.1%; Score 743; DB 1; Length 348;
Best Local Similarity 42.5%; Pred. No. 1.1e-43;
Matches 141; Conservative 63; Mismatches 126; Indels 2; Gaps 2;

QY 5 CHNINISGVKNNMNDVRASISLMLVLLITLVGNLIYVSIHFQKQHTPTMNLHS 64
DB 14 CYKNNESCTKTPYSPGPPSILYAVLFGAVLAAGNLMWIAIHPKQLHTPTPLIAS 73
QY 65 MATVDFLLGLVMPYSWVRSABHCWYFGEVFCIKHTSTDIMLSASIFHLSPISIDRYA 124
DB 74 LACADPLVGVVWMPSTYRSVSCWYFGDSYCKHTCPTSCFASLHLCISIDRYIA 133
QY 125 VCDPLRYAKNIIIVICWMIPIFSVSPAVPAFGMIFLEINFGAEIYYKHVCRGGSV 184
DB 134 VTDPLTYPTKFTVSYSICIVLWFFSVTSFSIYFTGANEGIBELVVA-LTCVGGCOA 192
QY 185 FFSKISGLVLTMTSPYIPGSIIMLCVYRRIYLAKEQARLISANOQKQIGLEMKGISQS 244
DB 193 PLNQWVLLICFL-L-FTIPVAVAVFLYSKIFLVAKIQARKISTVASQOSSSSYKERVAK 251
QY 245 KERKAVKTLGIWGVFLICWCPFFICTVMDPLHYIIPPTLNDVLIWFGYLNSTFNPVY 304
DB 252 RERKAKTLGLIAMAFLVSWLPYLDVAVIDAMNITTPYYVELVWCYVNSANPLIY 311
QY 305 APTFYWPKKALKMFLFKIFOKOSSRCKLPLE 336
DB 312 APTFYWPKKALKLIVSGKVLRTDSSTNLFSF 343

RESULT 13
TAR12_RAT 13
ID TAR12_RAT STANDARD; PRT; 333 AA.
AC Q923X8;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Trace amine receptor 12 (Iar-12).
GN Name=Iar12; Synonyms=Iar12; Tar12;
OS Rattus norvegicus (Rat).
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=21374364; PubMed=11459929; DOI=10.1073/pnas.151105198;
RA Borowicz B., Adam N., Jones K.A., Raddatz R., Artyushyn R.,
RA Ogozalek K.L., Durkin M.M., Lakhiani P.P., Bonini U.A., Pathirana S.,
RA Boyle N., Pu X., Kouranova E., Lichtblau H., Ochoa F.Y.,
RA Branchek T.A., Gerald C.;
RT "Trace amines: identification of a family of mammalian G protein-
RT coupled receptors.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:8966-8971(2001).
RL -1- FUNCTION: Orphan receptor. Could be a receptor for trace amines.
CC Trace amines are biogenic amines present in very low levels in
CC mammalian tissues. Although some trace amines have clearly defined
CC roles as neurotransmitters in invertebrates, the extent to which
CC they function as true neurotransmitters in vertebrates has
CC remained speculative. Trace amines are likely to be involved in a
CC variety of physiological functions that have yet to be fully
CC understood.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).

DR EMBL; AF380200; AAK1251.1; -.
DR RGD; 631392; Tar12.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR009132; Tracamine_recept.
DR Pfam; PF00001; 7tm1_1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PRINTS; PR01830; TRACAMINER.
DR PROSITE; PS00237; G_PROTEIN_RECEPTOR_F1_1.
DR PROSITE; PS0262; G_PROTEIN_RECEPTOR_F2_1.
KW G-protein coupled receptor; Glycoprotein; Multigene family;
KW Transmembrane.
FT DOMAIN 1 22 Extracellular (Potential).
FT TRAMSEM 23 43 1 (Potential).
FT DOMAIN 44 58 Cytoplasmic (Potential).
FT TRAMSEM 59 79 2 (Potential).
FT DOMAIN 80 99 Extracellular (Potential).
FT TRAMSEM 100 118 3 (Potential).
FT DOMAIN 119 141 Cytoplasmic (Potential).
FT TRAMSEM 142 162 4 (Potential).
FT DOMAIN 163 186 Extracellular (Potential).
FT TRAMSEM 187 207 5 (Potential).
FT DOMAIN 208 249 Cytoplasmic (Potential).
FT TRAMSEM 250 270 6 (Potential).
FT DOMAIN 271 283 Extracellular (Potential).
FT TRAMSEM 284 304 7 (Potential).
FT DOMAIN 305 333 Cytoplasmic (Potential).
FT DISULFID 95 180 By similarity.
FT CARBOHYD 9 N-linked (GlcNAc..?) (Potential).
SQ SEQUENCE 333 AA; 37467 MW; 65BD86131C1C95B CRC64;

Query Match 41.0%; Score 741; DB 1; Length 333;
Best Local Similarity 41.8%; Pred. No. 1.5e-43;
Matches 140; Conservative 70; Mismatches 123; Indels 2; Gaps 2;

QY 2 MPECHNINISGVKNNMNDVRASISLMLVLLITLVGNLIYVSIHFQKQHTPTMNL 61
DB 1 MQLCYEKLNRSVRSPPSPGRLILYAVFGAVLAAGNLMWIAIHPKQLHSFANFL 60
QY 62 IHSNATVDFLLGLVMPYSWVRSABHCWYFGEVFCIKHTSTDIMLSASIFHLSPISIDR 121
DB 61 VASLACADPLVGLVWMPSTYRSVSCWYFGDSYCKHTCPTSCFASLHLCISIDR 120
QY 122 YVAVCDPLRYAKNIIIVICWMIPIFSVSPAVPAFGMIFLEINFGAEIYYKHVCRG 181
DB 121 YVAVSDPLIYPTFTSVSGKCIPTFSLISIIYSFSLFYGVNAGAEDEL-VSALTCVGG 179
QY 182 CSVFEFSISGLVLTMTSPYIPGSIIMLCVYRRIYLAKEQARLISANOQKQIGLEMKG 241
DB 180 CQIADVNSWVFINL-L-FIVPALVMTVTSKIFLIAQQAONIERKQKQPARASSEYKDR 238
QY 242 SOSKERAIVKTLGIWGVFLICWCPFFICTVMDPLHYIIPPTLNDVLIWFGYLNSTFNP 301
DB 239 VAKRERKAATLGLIAMAFLVSWLPYLDVAVIDAMNITTPYYVELVWCYVNSANPLIY 298
QY 302 MTAFTFYWPKKALKMFLFKIFOKOSSRCKLPLE 336
DB 299 LIIAFTFYWPKKALKLIVTSGKILRENSATNLFSF 333

RESULT 14
TAR7_RAT 14
ID TAR7_RAT STANDARD; PRT; 344 AA.

CC	Q923Y3;		(Rel. 43, Created)
DT	29-MAR-2004		(Rel. 43, Last sequence update)
DT	29-MAR-2004		(Rel. 43, Last annotation update)
DT	25-OCT-2004		(Rel. 45, Last annotation update)
DE	Trace amine receptor 7 (Tar-7).		
GN	Name=Trar7; Synonyms=Tar7, Tar7;		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Vertebrate; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxId=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Sprague-Dawley;		
RX	MEDLINE=21374364; PubMed=11455929; DOI=10.1073/pnas.151105198;		
RA	Botwiner B., Adham N., Jones K.A., Raddatz R., Artyomshyn R.,		
RA	Ogozalek K.L., Durkin M.M., Lakhiani P.P., Bonini J.A., Pathirana S.,		
RA	Boyle N., Pu X., Koumanova E., Lichtblau H., Ochoa F.Y.,		
RA	Branchek T.A., Gerald C.;		
RT	"Trace amines: identification of a family of mammalian G protein-		
RT	coupled receptors."		
RL	Proc. Natl. Acad. Sci. U.S.A. 98:8966-8971(2001).		
CC	-I- FUNCTION: Orphan receptor. Could be a receptor for trace amines.		
CC	Trace amines are biogenic amines present in very low levels in		
CC	mammalian tissues. Although some trace amines have clearly defined		
CC	roles as neurotransmitters in invertebrates, the extent to which		
CC	they function as true neurotransmitters in vertebrates has		
CC	remained speculative. Trace amines are likely to be involved in a		
CC	variety of physiological functions that have yet to be fully		
CC	understood.		
CC	-I- SUBCELLULAR LOCATION: Integral membrane protein.		
CC	-I- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isdb-sib.ch/announce/		
CC	or send an email to license@isdb-sib.ch).		
CC	-----		
DR	EMBL; AF380195; AAK71246.1; --		
DR	RGD; 631386; Tar7		
DR	InterPro; IPR000276; GPCR_Rhodopsn		
DR	InterPro; IPR009132; Tracamine_recept.		
DR	Pfam; PF00001; 7tm_1; 1.		
DR	PRINTS; PR00237; GPCRRHODOPSIN.		
DR	PRINTS; PS01830; TRACAMINER.		
DR	PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.		
DR	PROSITE; PS50262; G_PROTEIN_RECP_F1_2; 1.		
KW	G-protein coupled receptor; Glycoprotein; Multigene family;		
KW	Transmembrane.		
FT	DOMAIN	1	33
FT	TRANSMEM	34	54
FT	DOMAIN	55	67
FT	TRANSMEM	68	88
FT	DOMAIN	89	102
FT	TRANSMEM	103	127
FT	DOMAIN	128	146
FT	TRANSMEM	147	167
FT	DOMAIN	168	196
FT	TRANSMEM	197	217
FT	DOMAIN	218	256
FT	TRANSMEM	257	277
FT	DOMAIN	278	295
FT	TRANSMEM	296	319
FT	DOMAIN	320	344
FT	DISULFID	104	189
FT	CARBOHYD	4	4
FT	CARBOHYD	18	18
FT	CARBOHYD	344 AA;	38043 MW;
SQ	SEQUENCE		
Query Match	40.6%;	Score 723;	DB 1; Length 344;
Best Local Similarity	41.2%;	Pred No. 5; 5e-43;	

	Matches	139,	Conservative	73,	Mismatches	121,	Indels	4;	Gaps	3,
Oy	2	MPFPHIINIISCVAKNMSNDVRASLYSLAMLIIITTVGNLIYVLSIHKQHLTPTNNML	61							
Dd	10	LQICENVVNNSCIKTTPSPSGRLVLIVWFGGALVACGMILVVISVLFHQHLSPPNFL	69							
Oy	62	IHSATVDPLGLCLMWPYSNWARSHEHCWGVEVCCKHTSTNDIMLSASIFHLFSIDR	121							
Dd	70	IASIASADPLVGISVMFPSSMWRSIESCMYFGDDFCSLSCDAAFCYSSLFHLCFISVDR	129							
Oy	122	YYAVCDLPLRYKAKNNILVICMTMFISWSVAVAFAGMIPLFLNEFKGEIEIYKKIHRCGG	181							
Dd	130	XIATEELVVPTRKTMSSVSGICISISWILLPLVVSANFYGISATGIENL-VSALNCVCG	188							
Oy	182	CSVFESKISGVLTFTMSFYIPGSIMLCVYRIYIYIAKEOARLI--SPANOKLQIGLEMKV	239							
Dd	189	COVANINDDWLVISFL-FPIPLTWMILIKSEIFPAKQAQVKIETSISSGSGESSLSHK	247							
Oy	240	GISSCKRKAVKTLTGIMGVFLWCPCPFICTWDPLAHYIIPRTLADVILMFWYLNSTF	299							
Dd	248	ARVARRERKAQKITGVTWAPFMWSMLPYTTDLTDLDAMGFITPAYVEICGMAYNSAM	307							
Oy	300	NPMVYAEFFYPWFRRKALKMMLFGKIFODSRCKFLFLE	336							
Dd	308	NPLIYAFFYPWFRRKAILKLISGKLTKGHSSRTSLFSE	344							

RESULT 15

ID	TAR4_RAT	STANDARD;	PRT;	345 AA.
AC	O923T5;			
DT	29-MAR-2004 (Rel. 43, Created)			
DT	29-MAR-2004 (Rel. 43, Last sequence update)			
DT	25-OCT-2004 (Rel. 45, Last annotation update)			
DE	Trace amine receptor 4 (Tar-4).			
CN	Name=Tar4; Synonyms=Ta4, Tar4;			
OS	Rattus norvegicus (Rat).			
OC	Eumetazoa; Chordata; Craniota; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Sprague-Dawley;			
EX	MEDLINE=21374364; PubMed=11459929; DOI=10.1073/pnas.151105198;			
RA	Borowsky B., Adam N., Jones K.A., Raddatz R., Artyomshyn R.,			
RA	Ogozalek K.L., Durkin M.M., Lakhani P.P., Bonini J.A., Pachirena S.,			
RA	Boyle N., Pu X., Kouzanova E., Iichblau H., Ochoa F.Y.,			
BR	Branchek T.A., Gerald C.;			
RT	"Trace amines: identification of a family of mammalian G protein-			
RT	coupled receptors.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 98:8966-8971(2001).			
CC	-1- FUNCTION: Orphan receptor. Could be a receptor for trace amines.			
CC	Trace amines are biogenic amines present in very low levels in			
CC	mammalian tissues. Although some trace amines have clearly defined			
CC	roles as neurotransmitters in invertebrates, the extent to which			
CC	they function as true neurotransmitters in vertebrates has			
CC	remained speculative. Trace amines are likely to be involved in a			
CC	variety of physiological functions that have yet to be fully			
CC	understood.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
DR	EMBL; AF380191; AAK71242.1; -.			
DR	RGD; 631384; Ta4			
DR	InterPro; IPR000276; GPCR_Rhodpsn.			

